











































- [10] Krishna R, Menzies T, Fu W. Too much automation? The bellwether effect and its implications for transfer learning. In: Proc. of the IEEE/ACM Int'l Conf. on Automated Software Engineering. 2016. 122–131.
- [11] Li ZQ, Jing XY, Zhu XK, Zhang HY. Heterogeneous defect prediction through multiple kernel learning and ensemble learning. In: Proc. of the IEEE Int'l Conf. on Software Maintenance and Evolution. 2017. 91–102.
- [12] Nam J, Pan SJ, Kim S. Transfer defect learning. In: Proc. of the 35th Int'l Conf. on Software Engineering. 2013. 382–391.
- [13] Peters F, Menzies T, Marcus A. Better cross company defect prediction. In: Proc. of the IEEE Working Conf. on Mining Software Repositories. 2013. 409–418.
- [14] Turhan B, Menzies T, Bener AB, *et al.* On the relative value of cross-company and within-company data for defect prediction. *Empirical Software Engineering*, 2009,14(5):540–578.
- [15] Zimmermann T, Nagappan N, Gall H, Giger E, Murphy B. Cross-project defect prediction: A large scale experiment on data vs. domain vs. process. In: Proc. of the Joint Meeting of the European Software Engineering Conf. and the ACM SIGSOFT Symp. on the Foundations of Software Engineering. 2009. 91–100.
- [16] Liu WS, Chen X, Gu Q, Liu SL, Chen DX. A noise tolerable feature selection framework for software defect prediction. *Chinese Journal of Computers*, 2018,41(3):506–520 (in Chinese with English abstract).
- [17] Liu WS, Liu SL, Gu Q, Chen JQ, Chen X, Chen DX. Empirical studies of a two-stage data preprocessing approach for software fault prediction. *IEEE Trans. on Reliability*, 2016,65(1):38–53.
- [18] Chen X, Zhao YQ, Wang QP, Yuan ZD. Multi: Multi-objective effort-aware just-in-time software defect prediction. *Information and Software Technology*, 2018,93:1–13.
- [19] Chen X, Zhang D, Zhao YQ, Cui ZQ, Ni C. Software defect number prediction: Unsupervised vs supervised methods. *Information and Software Technology*, 2019,106:161–181.
- [20] Liu WS, Chen X, Gu Q, Liu SL, Chen DX. A cluster analysis based feature selection method for software defect prediction. *Scientia Sinica Informationis*, 2016,46(9):1298–1320 (in Chinese with English abstract).
- [21] He JY. Search based semi-supervised ensemble learning research for cross-project defect prediction [MS. Thesis]. Tianjin: Tianjin University, 2017.
- [22] Ghotra B, McIntosh S, Hassan AE. Revisiting the impact of classification techniques on the performance of defect prediction models. In: Proc. of the Int'l Conf. on Software Engineering. 2015. 789–800.
- [23] Peters F, Menzies T, Layman L. Lace2: Better privacy-preserving data sharing for cross project defect prediction. In: Proc. of the Int'l Conf. on Software Engineering. 2015. 801–811.
- [24] Tantithamthavorn C, McIntosh S, Hassan AE, Ihara A, Matsumoto K. The impact of mislabelling on the performance and interpretation of defect prediction models. In: Proc. of the Int'l Conf. on Software Engineering. 2015. 812–823.
- [25] Jing XY, Wu F, Dong XW, Qi FM, Xu BW. Heterogeneous cross-company defect prediction by unified metric representation and CCA-based transfer learning. In: Proc. of the Joint Meeting on Foundations of Software Engineering. 2015. 496–507.
- [26] Kim MJ, Nam JC, Yeon JY, Choi SW, Kim SH. Remi: Defect prediction for efficient API testing. In: Proc. of the Joint Meeting on Foundations of Software Engineering (ESEC/FSE 2015). 2015. 990–993.
- [27] Nam JC, Kim SH. Clami: Defect prediction on unlabeled datasets (t). In: Proc. of the Int'l Conf. on Automated Software Engineering. 2015. 452–463.
- [28] Radjenović D, Heričko M, Torkar R, *et al.* Software fault prediction metrics: A systematic literature review. *Information & Software Technology*, 2013,55(8):1397–1418.
- [29] Menzies T, Greenwald J, Frank A. Data mining static code attributes to learn defect predictors. *IEEE Trans. on Software Engineering*, 2007,33(1):2–13.
- [30] Song QB, Jia ZH, Shepperd M, Ying S. A general software defect-proneness prediction framework. *IEEE Trans. on Software Engineering*, 2011,37(3):356–370.
- [31] Agrawal A, Menzies T. Is “better data” better than “better data miners”? On the benefits of tuning smote for defect prediction. In: Proc. of the Int'l Conf. on Software Engineering. 2018. 1050–1061.
- [32] Yu X, Liu J, Yang ZJ, Jia XY, Ling Q, Ye SZ. Learning from imbalanced data for predicting the number of software defects. In: Proc. of the Int'l Symp. on Software Reliability Engineering. 2017. 78–89.
- [33] Xu Z, Liu J, Yang ZJ, An GG, Jia XY. The impact of feature selection on defect prediction performance: An empirical comparison. In: Proc. of the Int'l Symp. on Software Reliability Engineering. 2016. 309–320.
- [34] Fukushima T, Kamei Y, McIntosh S, Yamashita K, Ubayashi N. An empirical study of just-in-time defect prediction using cross-project models. In: Proc. of the 11th Working Conf. on Mining Software Repositories. 2014. 172–181.

- [35] He JY, Meng ZP, Chen X, Wang Z, Fan XY. Semi-supervised ensemble learning approach for cross-project defect prediction. *Ruan Jian Xue Bao/Journal of Software*, 2017,28(6):1455–1473 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5228.htm> [doi: 10.13328/j.cnki.jos.005228]
- [36] Ma Y, Luo GC, Zeng X, Chen AG. Transfer learning for cross-company software defect prediction. *Information and Software Technology*, 2012,54(3):248–256.
- [37] Wang S, Liu TY, Tan L. Automatically learning semantic features for defect prediction. In: *Proc. of the Int'l Conf. on Software Engineering*. 2016. 297–308.
- [38] Chen L, Fang B, Shang ZW, Tang YY. Negative samples reduction in cross-company software defects prediction. *Information and Software Technology*, 2015,62:67–77.
- [39] He P, Li B, Ma YT. Towards cross-project defect prediction with imbalanced feature sets. arXiv preprint arXiv:1411.4228, 2014.
- [40] Nam JC, Kim SH. Heterogeneous defect prediction. In: *Proc. of the Joint Meeting of the European Software Engineering Conf. and the ACM SIGSOFT Symp. on the Foundations of Software Engineering*. 2015. 508–519.
- [41] Zhong S, Khoshgoftaar TM, Seliya N. Unsupervised learning for expert-based software quality estimation. In: *Proc. of the 2004 8th IEEE Int'l Symp. on High Assurance Systems Engineering*. 2004. 149–155.
- [42] Zhang F, Zheng Q, Zou Y, Hassan AE. Cross-project defect prediction using a connectivity-based unsupervised classifier. In: *Proc. of the Int'l Conf. on Software Engineering*. 2016. 309–320.
- [43] Yang YB, Zhou YM, Liu JP, Zhao YY, Lu HM, Xu L, Xu BW, Leung H. Effort-aware just-in-time defect prediction: Simple unsupervised models could be better than supervised models. In: *Proc. of the 24th ACM SIGSOFT Int'l Symp. on Foundations of Software Engineering*. 2016. 157–168.
- [44] Zhou YM, Yang YB, Lu HM, Chen L, Li YH, Zhao YY, Qian JY, Xu BW. How far we have progressed in the journey? An examination of cross-project defect prediction. *ACM Trans. on Software Engineering and Methodology*, 2018,27(1):Article No.1.
- [45] Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans. on Knowledge & Data Engineering*, 2010,22(10):1345–1359.
- [46] Zhuang FZ, Luo P, Xiong H, Xiong YH, He Q, Shi ZZ. Cross-domain learning from multiple sources: A consensus regularization perspective. *IEEE Trans. on Knowledge & Data Engineering*, 2010,22(12):1664–1678.
- [47] Dai WY, Yang Q, Xue GR, Yu Y. Boosting for transfer learning. In: *Proc. of the 24th Int'l Conf. on Machine Learning*. 2007. 193–200.
- [48] Dai WY, Xue GR, Yang Q, Yu Y. Transferring naive Bayes classifiers for text classification. In: *Proc. of the National Conf. on Artificial Intelligence*. 2007. 540–545.
- [49] Swarup S, Ray SR. Cross-domain knowledge transfer using structured representations. In: *Proc. of the National Conf. on Artificial Intelligence*. 2006. 506–511.
- [50] Ni C. Research on software defect prediction based on transfer learning [MS. Thesis]. Nanjing: Nanjing University, 2017.
- [51] Wu Q. Cross-project defect prediction based on transfer learning [MS. Thesis]. Changchun: Jilin University, 2018.
- [52] Yu L, Liu H. Efficient feature selection via analysis of relevance and redundancy. *Journal of Machine Learning Research*, 2004, 5(12):1205–1224.
- [53] Kira K, Rendell LA. The feature selection problem: Traditional methods and a new algorithm. In: *Proc. of the 10th National Conf. on Artificial Intelligence*. 1992. 129–134.
- [54] D'Ambros M, Lanza M, Robbes R. Evaluating defect prediction approaches: A benchmark and an extensive comparison. *Empirical Software Engineering*, 2012,17(4):531–577.
- [55] Peters F, Menzies T. Privacy and utility for defect prediction: Experiments with MORPH. In: *Proc. of the Int'l Conf. on Software Engineering*. 2012. 189–199.
- [56] Wu RX, Zhang HY, Kim SH, Cheung SC. Relink: Recovering links between bugs and changes. In: *Proc. of the ACM Sigsoft Symp. and the European Conf. on Foundations of Software Engineering*. 2011. 15–25.
- [57] D'Ambros M, Lanza M, Robbes R. An extensive comparison of bug prediction approaches. In: *Proc. of the Mining Software Repositories*. 2010. 31–41.
- [58] Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1945,1(6):80–83.
- [59] Janez Ar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 2006,7(1):1–30.
- [60] Liu SL, Chen X, Liu WS, Chen JQ, Gu Q, Chen DX. Fecar: A feature selection framework for software defect prediction. In: *Proc. of the Computer Software and Applications Conf*. 2014. 426–435.
- [61] Gao KH, Khoshgoftaar TM, Wang HJ, Seliya N. Choosing software metrics for defect prediction: An investigation on feature selection techniques. *Software Practice & Experience*, 2011,41(5):579–606.

- [62] Kim SH, Zhang HY, Wu RX, Gong L. Dealing with noise in defect prediction. In: Proc. of the Int'l Conf. on Software Engineering. 2011. 481–490.
- [63] Herbold S. CrossPare: A tool for benchmarking cross-project defect predictions. In: Proc. of the Int'l Conf. on Automated Software Engineering Workshop. 2015. 90–96.
- [64] He ZM, Shu FD, Yang Y, Li MS, Wang Q. An investigation on the feasibility of cross-project defect prediction. Automated Software Engineering, 2012,19(2):167–199.
- [65] Rahman F, Posnett D, Devanbu P. Recalling the “imprecision” of cross-project defect prediction. In: Proc. of the ACM SIGSOFT Symp. on the Foundations of Software Engineering. 2012. 1–11.
- [66] Fan LL, Su T, Chen S, Meng GZ, Liu Y, Xu LH, Pu GG. Efficiently manifesting asynchronous programming errors in android apps. In: Proc. of the 33rd ACM/IEEE Int'l Conf. on Automated Software Engineering. 2018. 486–497.
- [67] Fan LL, Su T, Chen S, Meng GZ, Liu Y, Xu LH, Pu GG, Su ZD. Large-scale analysis of framework-specific exceptions in android apps. In: Proc. of the 40th Int'l Conf. on Software Engineering. 2018. 408–419.
- [68] Su T, Meng GZ, Chen YT, Wu K, Yang WM, Yao Y, Pu GG, Liu Y, Su ZD. Guided, stochastic model-based GUI testing of android apps. In: Proc. of the 2017 11th Joint Meeting on Foundations of Software Engineering. 2017. 245–256.
- [69] Lewis C, Lin ZP, Sadowski C, Zhu XY, Ou R, Whitehead EJ. Does bug prediction support human developers? Findings from a google case study. In: Proc. of the 2013 Int'l Conf. on Software Engineering. 2013. 372–381.

#### 附中文参考文献:

- [1] 陈翔,顾庆,刘望舒,刘树龙,倪超.静态软件缺陷预测方法研究.软件学报,2016,27(1):1–25. <http://www.jos.org.cn/1000-9825/4923.htm> [doi: 10.13328/j.cnki.jos.004923]
- [2] 王青,伍书剑,李明树.软件缺陷预测技术.软件学报,2008,19(7):1565–1580. <http://www.jos.org.cn/1000-9825/19/1565.htm> [doi: 10.3724/SP.J.1001.2008.01565]
- [5] 陈翔,王莉萍,顾庆,王赞,倪超,刘望舒,王秋萍.跨项目软件缺陷预测方法研究综述.计算机学报,2018,41(1):254–274.
- [16] 刘望舒,陈翔,顾庆,刘树龙,陈道蓄.一种面向软件缺陷预测的可容忍噪声的特征选择框架.计算机学报,2018,41(3):506–520.
- [20] 刘望舒,陈翔,顾庆,刘树龙,陈道蓄.软件缺陷预测中基于聚类分析的特征选择方法.中国科学:信息科学,2016,46(9):1298–1320.
- [21] 何吉元.基于搜索的半监督集成跨项目软件缺陷预测方法研究[硕士学位论文].天津:天津大学,2017.
- [35] 何吉元,孟昭鹏,陈翔,王赞,樊向宇.一种半监督集成跨项目软件缺陷预测方法.软件学报,2017,28(6):1455–1473. <http://www.jos.org.cn/1000-9825/5228.htm> [doi: 10.13328/j.cnki.jos.005228]
- [50] 倪超.基于迁移学习的软件缺陷预测方法研究[硕士学位论文].南京:南京大学,2017.
- [51] 吴琦.基于迁移学习的跨项目软件缺陷预测[硕士学位论文].长春:吉林大学,2018.



倪超(1990—),男,江苏南京人,博士生,主要研究领域为软件缺陷预测.



顾庆(1972—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为软件质量保障,分布式计算.



陈翔(1980—),男,博士,副教授,CCF 高级会员,主要研究领域为软件缺陷预测,软件缺陷定位,回归测试和组合测试.



黄启国(1979—),男,博士生,主要研究领域为软件缺陷预测.



刘望舒(1987—),男,博士,讲师,CCF 专业会员,主要研究领域为软件缺陷预测,软件缺陷定位.



李娜(1991—),女,博士生,主要研究领域为机器学习.