中国科学院软件研究所学术年会’2019
暨计算机科学国家重点实验室开放周

学术论文

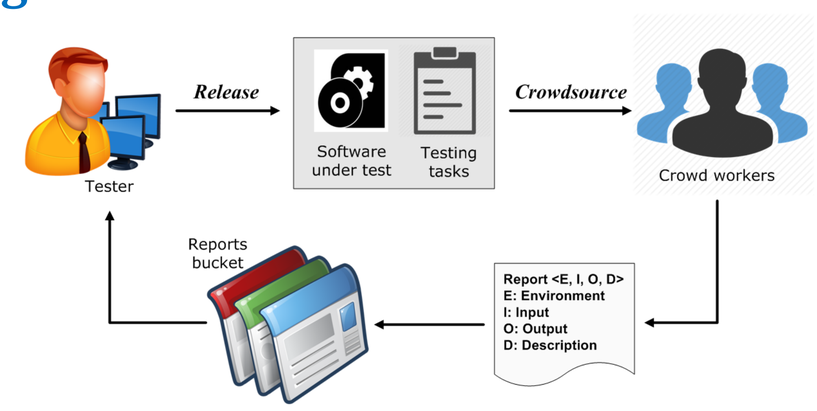
iSENSE: Completion-Aware Crowdtesting Management
任务完成感知的众测管理方法

Junjie Wang, Ye Yang, Rahul Krishna, Tim Menzies, Qing Wang

In Proceedings of the 41st ACM/IEEE International Conference on Software Engineering (ICSE 2019)
ICSE 2019 ACM SIGSOFT Distinguished Paper award

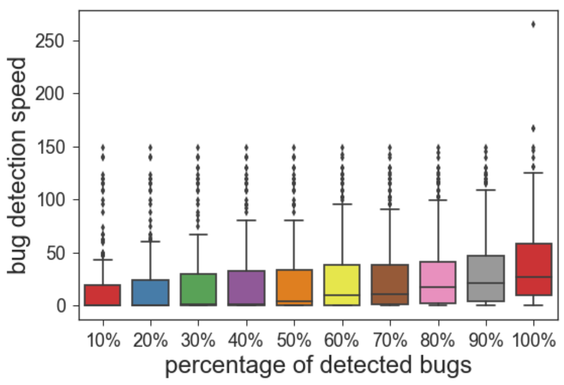
联系人：王青 王俊杰　　联系方式：{wq, junjie}@iscas.ac.cn

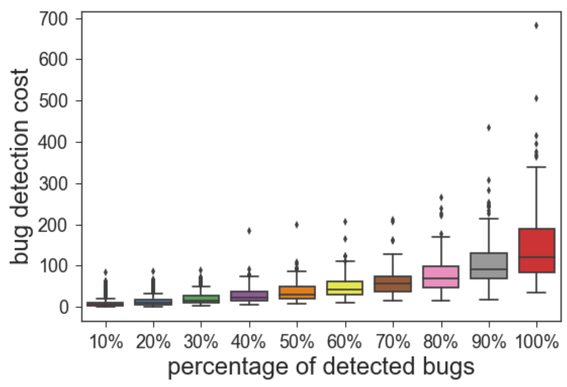
Background



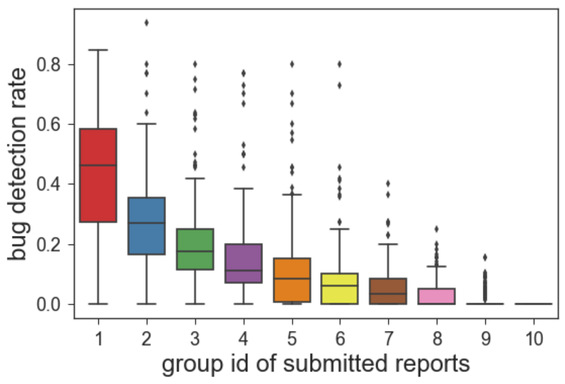
- Crowdtesting entrusts tasks to online crowdworkers whose diverse testing environments, background, and skill sets could significantly contribute to more reliable, cost-effective, and efficient testing results.
- Trade-offs such as “how much testing is enough” are critical yet challenging project decisions.
- Current practices usually set up either a fixed period (e.g., 5 days) or a fixed number of participant (e.g., recruiting 400 crowd workers) for the close criteria.

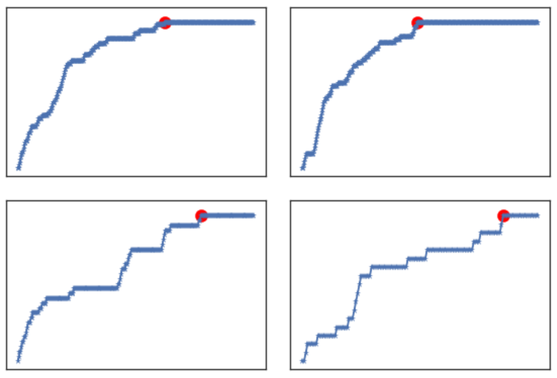
Observations From A Pilot Study about bug arrival patterns of crowdtesting





Finding 1: Large Variation in Bug Detection Speed and Cost





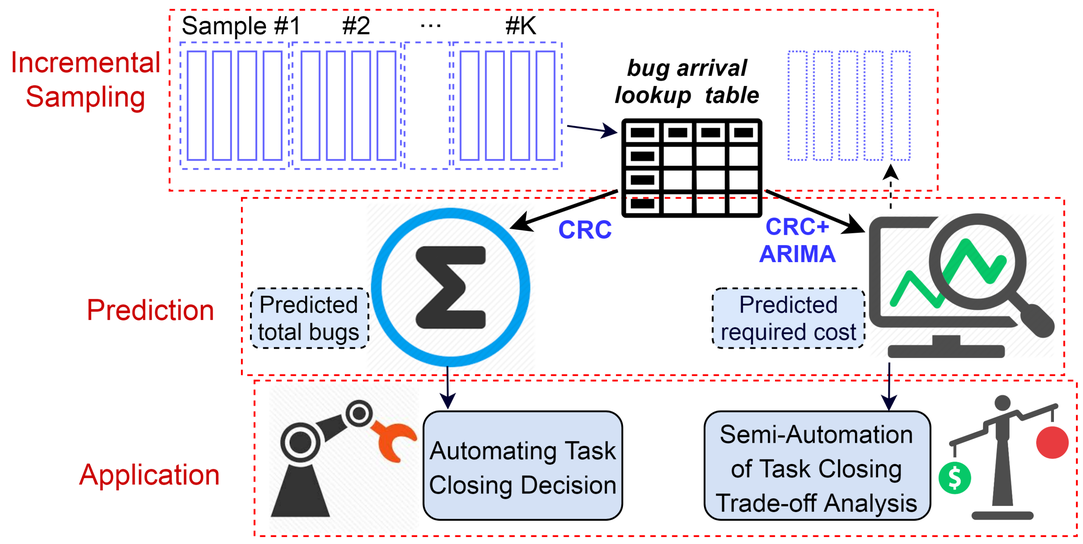
Finding 2: Decreasing Bug Detection Rates Over Time

Finding 3: Plateau Effect of Bug Arrival Curve

Current decision making is largely done by guesswork.
This results in **low cost-effectiveness of crowdtesting**.
A more effective alternative would be to dynamically monitor the crowdtesting process and provide actionable decision support for task closing to save unnecessary cost wasting on later arriving reports.

Approach

Propose completion-aware crowdtesting management approach iSENSE to **raise the awareness of testing progress and facilitate decision making.**



- Adopt an incremental sampling process to model crowdtesting reports. Convert the raw crowdtesting reports arrived chronologically into groups and generates a bug arrival lookup table to characterize the bug arrival information, i.e., bug and duplicate information.
- Integrate two models, i.e. Capture-ReCapture models and Autoregressive Integrated Moving Average model, to predict 1) the total number of bugs contained in the software, and 2) the required cost for achieving certain test objectives, respectively.
- Apply such estimates to support two typical crowdtesting decision scenarios, i.e., automating task closing decision, and semi-automation of task closing trade-off analysis.

Experiment

- 218 mobile application testing tasks with 46434 submitted reports from Baidu crowdtesting platform.
- MRE of prediction (on total bugs, and required cost) are both below 6%, with about 10% standard deviation.
- The automation of task closing can make crowdtesting more cost-effective, i.e., **a median of 100% bugs can be detected with 30% saved cost.**
- ISENSE provides practical insights to help managers make trade-off analysis on which task to close or when to close.