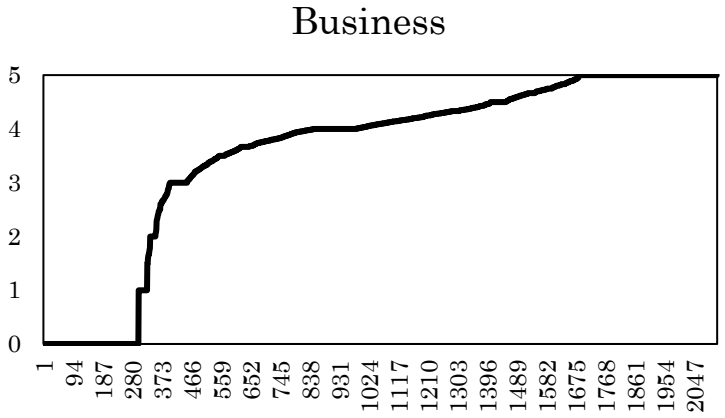
Threshold Selection in SAFER

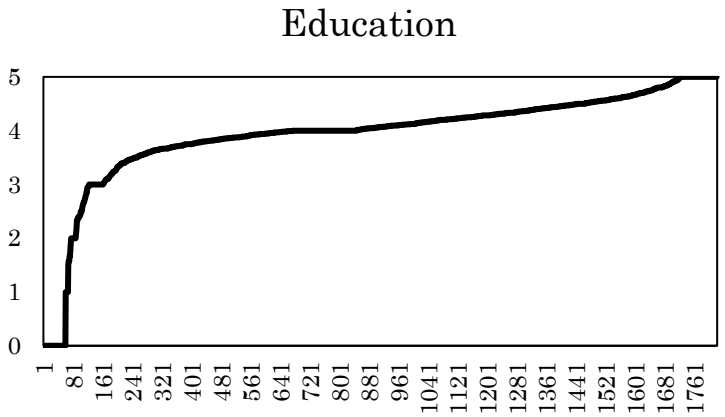
HE JIANG, Dalian University of Technology
JINGXUAN ZHANG, Nanjing University of Aeronautics and Astronautics
XIAOCHEN LI, Dalian University of Technology
ZHILEI REN, Dalian University of Technology
DAVID LO, Singapore Management University
XINDONG WU, University of Vermont
ZHONGXUAN LUO, Dalian University of Technology

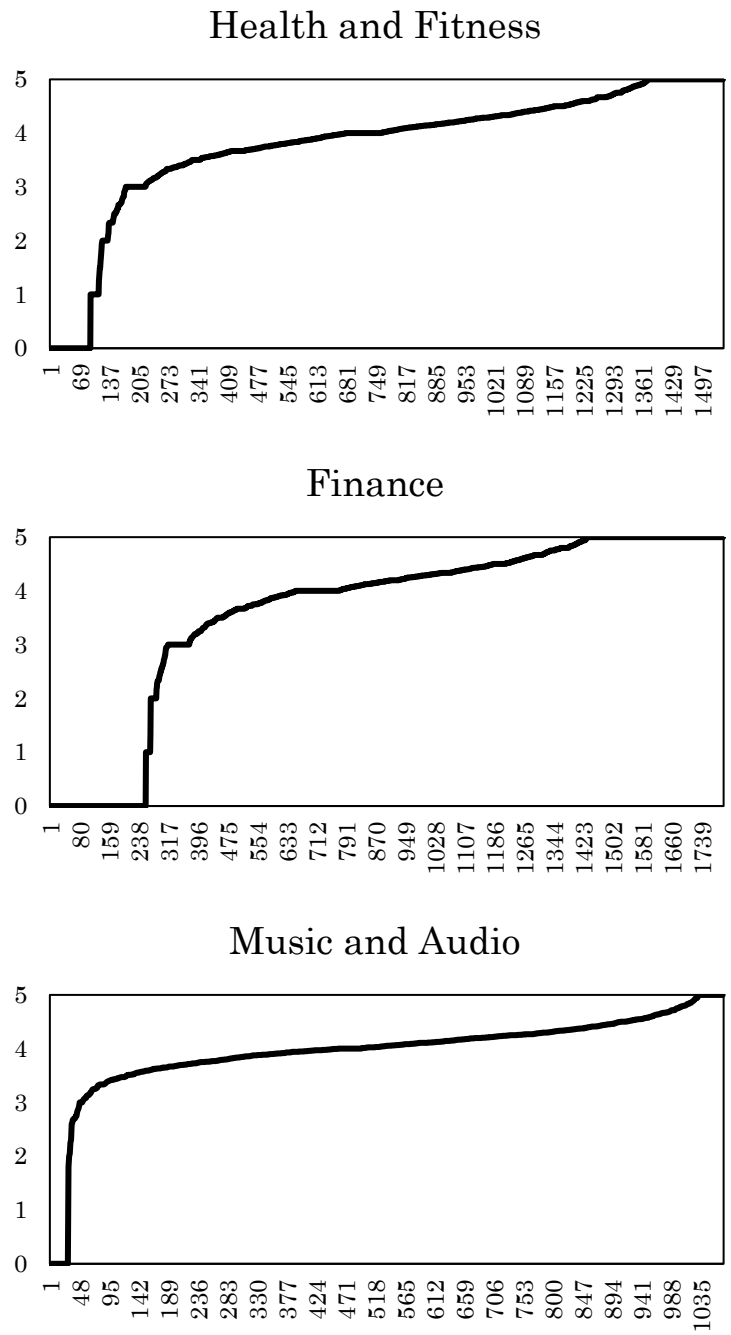
In the paper, we make some choices to some parameters, and set them to predefined values, namely the average rating in Reference App Filter, Rating Count in Reference App Filter, and Topic Number in Topic Model. This technical report tries to explore the effectiveness of the parameter selections.

1. Average Rating in Reference App Filter

To get inspired from high quality similar apps, SAFER needs to filter out apps with low quality by setting up the average rating threshold. The average rating in Reference App Filter is set up to 3 in the paper. For the apps in the Reference App Repository, we obtain their user ratings for each category, and rank them in descending order. The following figures show the distributions of user ratings in the five categories.





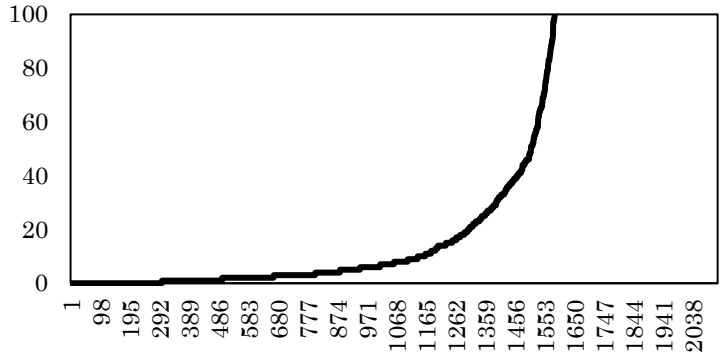


From the figures we can see that, user scores of a large scale of apps are more than 3, and only a fraction of use scores are less than 3. Setting the user rating to 3 is reasonable to retain the high quality competing apps.

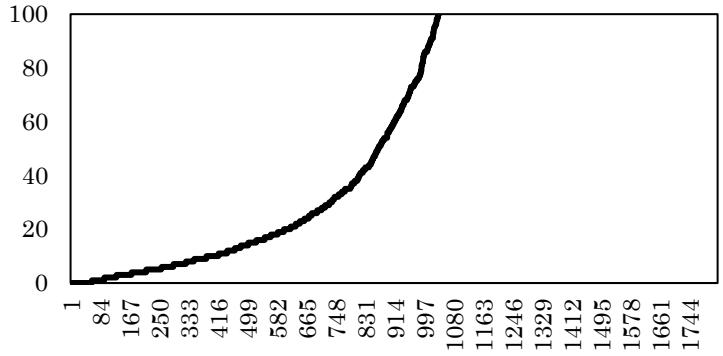
2. Rating Count in Reference App Filter

The same as the user rating, we also introduce the rating count to filter out low quality apps to reduce the bias of only introducing user rating. We also obtain the rating count and rank them in descending order for the five categories, and the following figures show the results.

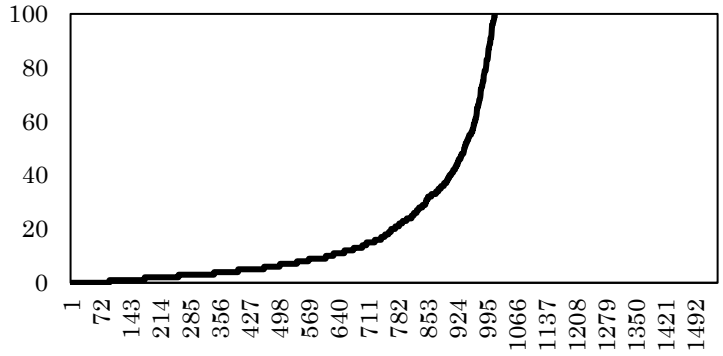
Business



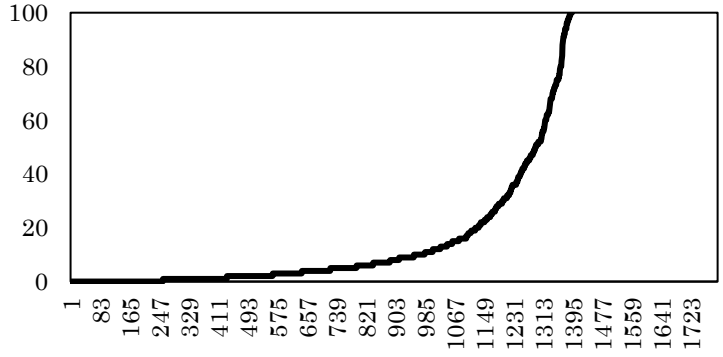
Education

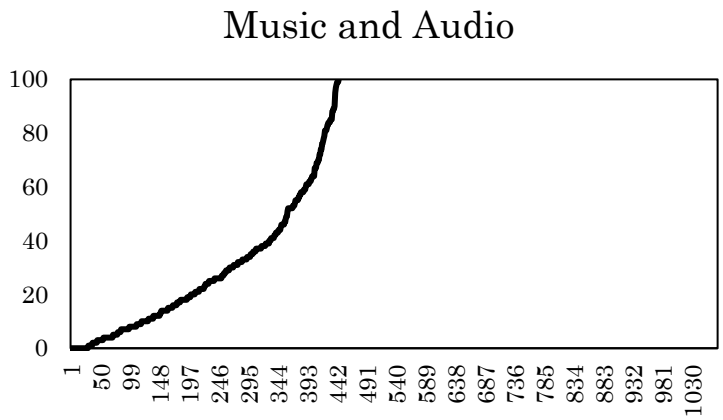


Health and Fitness



Finance



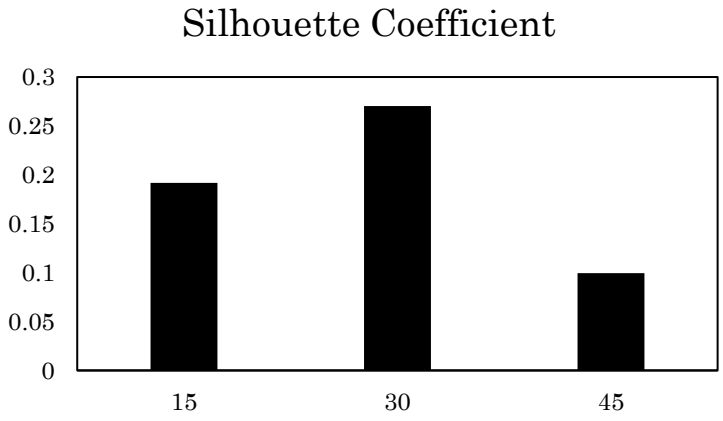


For each category, more than 500 high quality reference apps will be retained, which is large enough to perform domain analysis.

3. Topic Number in Topic Model

The topic number influences the performance of topic model. In the paper, we set it to the default value in TMT, i.e., 30. To test whether it is a good value, we select one category as a case study, namely *Music and Audio*. We use the same method proposed by Panichella et al., which considers topic model results as clusters and introduces Silhouette Coefficient (ranges from -1 to 1) to evaluate the qualities of topic model [Panichella et al. 2013]. We select anther two topic number values (15 and 45) to compare against the default value (30), to show whether the default value is a good value.

The following figure shows the Silhouette Coefficient for the topic numbers of 15, 30, and 45. We can see that all of them are positive values. Among them, topic number setting to 30 achieves the best results. That is to say, the default value of 30 approximates the optimal value.



[Annibale Panichella, Bogdan Dit, Rocco Oliveto, Massimiliano Di Penta, Denys Poshyvanyk, and Andrea De Lucia. 2013. How to Effectively Use Topic Models for Software Engineering Tasks? An Approach based on Genetic Algorithms. In Proceedings of the International Conference on Software Engineering (ICSE'13). pp.522-531.]