

《软件需求工程》实验二 实验报告

一、小组成员信息：

161220001 艾山江·吐尔逊

161220138 吴晗

161220179 周科

171860522 沈天琪

二、成绩分配比例：

每人各 25%

三、实验简介：

选定一个开源 IDE 项目，确定可能的信息来源，获取有效信息，对所获取的需求进行优先级排序。

四、实验过程：

本次实验延续实验一的实验成果完成。实验一用了两种方式来确定需求：一是通过爬取数据进行词频统计来获取，二是通过主页介绍来倒推获取。对于需求进行优先级排序，我们进行了 4 人对每项的独立打分。借由 1-5 分的打分结果，取 4 人均分来分定“优”“中”“差”三个等级，而后根据等级进行优先级排序。

而对于这次的效果验证，由于实验一的需求获取来源于两个方面，我们也分成两种方式来验证。

对于词频统计获取到的需求，我们根据开发者对每个 issue 的管理标签来确定其等级。例如打上“important”标签的 issue 在开发者眼里就是优先级最高的一类，而一旦我们根据词频得到的某个需求在这个 label 文件内被提及，我们便认定，此需求的预估开发者评级为“优”。相应的，在“invalid”标签下被涉及到的需求，我们认定其预估开发者评级为“差”。

对于官网主页倒推得到的需求结果，我们分为两个部分。第一个是根据官网 docs 部分获取的需求，由于这一栏都是对此 IDE 的功能和特色的介绍，我们可以认为主要的功能对标的就是最优先的需求，所以我们把这类总结出来的需求的预估开发者评级评定为优。第二个是根据官网 updates 部分获取的需求，这部分需求主要由软件的更新和修复日志得出。我们根据修复的快慢程度和被提及频率，以及功能更新的先后顺序，来作为此类需求的预估开发者评级。

我们根据 4 人打分的结果来作为软件需求的优先级排序结果。以后续分别得出的预估开发者评级作为对照结果，对我们的评级方法的效果进行分析。

五、实验任务分配：（具体完成内容见各自文件夹目录下的实验报告）

艾山江·吐尔逊：检索代码的编写，合作完成对由词频统计获得的需求的优先级评定（代码部分）。

吴晗：对由官网的 updates 部分获取的需求进行优先级评定。

周科：总实验报告的整合与撰写，合作完成对由词频统计获得的需求的优先级评定（人工分析部分）。

沈天琪：编写爬虫程序爬取 VSCode 的 issue 信息和 label 信息交由后续处理；对由官网的 docs 部分获取的需求进行优先级评定

六、实验结果展示：

实验结果存在于本实验报告同目录下的 excel 文件，名为“实验结果.xlsx”。

第一列为需求分类项。

第二列为具体的需求。其中用颜色对其进行了划分：

黄色为官网主页 docs 部分获取的需求。

绿色为官网主页 updates 部分获取的需求。

橙色为 docs 以及 updates 部分均有涉及的部分。

蓝色为 updates 中未被再次提及的部分。

粉色为词频统计得出的需求部分。

第三四五六列分别是沈天琪、吴晗、周科、艾山江各自的打分情况。

第七列为四人均分。

第八列为实际的评级结果（优、中、差三类，其中优为最高优先级，依次递减）。

第九列为预估开发者评级，作为实验结果对照。

七、方法效果分析：

综合来看，以程序员作为软件使用者的视角来看待需求的优先情况，和软件开发者实际的需求优先级认知效果，还是很相近的。排序结果大致吻合，偏差较大的出现在一些有特定术语存在的需求上，我们四人可能未能实际接触过此类功能，例如“Emmet 支持”、“HTMLARIA 属性引用链接”等。而最贴切的是对于“优”类的评定，此类一般涉及到“智能”、“自动”、“简化便捷”等词汇，可见的是程序员对于 SmartIDE 的 smart 需求越来越强烈，软件更加智能更加易用是需求，也是趋势。