



自动分析软件缺陷报告间相关性的方法研究

<http://www.firstlight.cn> 2010-06-01

针对缺陷报告相关性分析的研究主要采用计算其文本信息相似度的方法使其查全率和查准率并不理想，提出了一种将结构化信息相似度与文本信息相似度计算相结合的方法，即同时提取出缺陷报告中的文本信息（包括主题和详细描述）以及结构化信息（包括补丁、异常堆栈和代码片段），从缺陷外部表现和内部特征两个角度共同衡量缺陷报告间的相关性。通过对Eclipse系统中的1000个缺陷报告进行实验，结果显示，增加结构化信息相似度计算，可以有效地将缺陷报告间相关性分析的查准率和查全率均提高到90%左右。

[存档文本](#)