

杰出青年人才发展专项计划年度进展报告书

(2010 年度)

姓名	杨叶	课题名称	经验软件工程方法研究
资助金额	160 万	课题起止时间	2009. 9–2013. 9
资助类别	应用基础类	所在部门	基础软件国家工程研究中心
研究工作主要进展和阶段性成果（含论文、研究生培养等）			
本年度的主要工作围绕两方面开展： 1，基于多版本、多类型的软件演化数据，开展缺陷预测、成本估算和风险管理等方法和技术的经验研究，支持对软件演化的缺陷预测工具选取、成本估算决策分析以及风险管理策略等方面的经验提取和共享； 2，基于淘宝网、人人网等社区软件的演化数据，开展对软件生态群落演化的预测、评估和分析技术的经验研究，支持对软件开发组织、软件服务运营商预测、分析和管理其需求演化，实现软件生态系统的持续健康的演化和发展； 3，结合 Semantic Web 技术提出了新的经验库构建技术，支持对经验和数据的标注、检索和分析，并开发原型工具。 下面详细介绍主要的工作进展： 1. 缺陷预测方法对比分析 提出了相应的缺陷分类预测对比研究方法。该方法考虑确定实验目标、采集实验数据、分析数据集特征、构造缺陷分类预测模型、确定预测模型评价方法、假设检验以及有效性评估七个方面，通过规范研究过程和方法，以提高经验研究实验的可重复性和结果的可靠性，以及不同实验结果间的可对比性。最后，使用该方法，基于来自开源和闭源软件的数据集，给出了软件缺陷分类预测的研究实例，从分类器、数据预处理技术、所研究的软件项目开发类型、以及数据集使用的度量元统计粒度四个方面进行预测性能的对比分析，验证了所提出对比研究方法的可行性并给出了相应的经验性结论。目前正在进一步的研究中。 2. 基于用例和经验库的软件规模和成本估算方法 用例作为一种捕获和描述用户需求的客观表现形式，越来越多的作为度量软件规模和软件总工作量指标使用。在基于用例的重用估算研究中我们着眼于用例的特性，提出了两种基于用例的软件成本估算模型。一种方法采用用例点估算模型和 COCOMO II 相结合的组合估算方法，以用例规约文档为标准，改进用例点估算方法，估算软件规模以及软件成本。另一种方法则通过对重用开发中用例的分类和权值的分配，基于组织内的历史数据，通过回归方式建立重用估算模型，进行工作量估算。在基于用例的估算研究上，我们研究并开发了相应的估算工具。该工具已集成到软件过程管理平台产品 QONE 上。目前正在进一步的研究中。			

3. 基于人员能力与项目经验的进度风险风险分析方法的研究

软件风险指软件过程和产品可能遭受的影响和损失，是软件与生俱来的特质，并与软件的成本和进度紧密关联。课题组一直专注于风险管理的研究，包含风险识别、分析、评估等关键技术。在本课题的资助下，我们通过对软件开发过程的模拟仿真，定量的分析项目中人员能力和经验对项目进度造成的影响，并通过建立组织能力、人力资源和项目计划三个数据模型，建立项目调度和任务执行的仿真过程，对项目开发过程进行仿真，并通过 SCEDRisk 工具实现进度风险的自动化分析和图形化显示。目前正在进一步的研究中。

4. 软件生态群落演化的预测、评估和分析技术

软件的演化要求即时响应用户需求，因此这种演化是频繁发生、敏捷响应的。特别是一些服务于大众的网络化软件拥有大规模的用户，这些用户的用户偏好和使用模式既有共性，也有不同。其演化是在群体智能的驱动下进行的，并体现用户偏好。同时演化过程中的数据（如演化频率、需求和服务的匹配程度、服务冲突记忆、演化的行为模式等）可以为软件服务运营商提供知识，使他们可以更好地理解客户需求，预测演化的未来方向。我们提出了基于用户群体智能和用户保留率的演化预测方法，支持大众交互模式下的软件演化和服务组合改进及其评估。目前正在进一步的研究中。

5. 经验库构建技术

逐步明确和完善软件经验库的构建方案并已完成原型。经验库建设的短期目标包括：

- 以一种通用的格式上传经验数据，通过解析，将数据存储到经验库中；
- 用户可在经验库中浏览，查找，筛选和定制数据；
- 以 Web 方式实现经验库，并强调其易用性；

长期目标为不同学科背景研究人员共同交流和分享软件工程经验研究的平台，实现各学科与软件工程的交叉，具体包括：

- 自动分析和统计数据和经验，并实现数据和经验推送，个性化和定制化。
- 建立用户研究兴趣小组，对不同的研究小组分配不同的数据和经验，实现订阅功能。
- 结合数据挖掘和统计机器学习算法自动分析数据和经验并采用 Semantic Web 帮组用户理解从经验库中获取的经验。

目前正在进一步的研究和开发中。

6. 国际合作

2010 年 3 月到 4 月，美国西维吉尼亚大学、南加州大学系统与软件工程中心的两名博士生来到软件所本课题团队分别交流了 2 个月，在本课题负责人的指导下完成了关于估算数据收集和分析的一些合作研究，并撰写了相关的学术论文，已发表于国际学术会议。

2010 年 5 月，邀请美国 MIT 的 Ricardo Valerdi 副教授来访，开展软件成本估算方面的合作；2010 年 7 月 7-10 日，软件所参与主办的软件过程国际会议 ICSP2009(International Conference on Software Process)在德国 Paderborn 举行。会议由德国 Paderborn 大学承办，本课题负责人担任此次会议的程序委员会联合主席，会议吸引了来自学术界和产业界的 50 余名参与者，就“过程模型”、“过程仿真建模”、“过程变更管理”、和“过程分析和度量”等多个专题进行了广泛的讨论。

2010 年 7-9 月，回访美国 MIT 的 Ricardo Valerdi 副教授，开展软件成本估算方面的合作；2010 年 9 月，本课题负责人参加 ISERN2010 会议，围绕本课题进展及问题组织了一次相关的讨论会，得到了一些较有针对性的反馈意见，在目前的工作中正逐步实施；

2010 年 10 月, 邀请美国 RPI 的丁力研究员来访, 开展 Semantic Web 方面的合作。
2010 年 11 月 2 日至 5 日, 2010 软件成本建模与 COCOMO 国际论坛在美国南加州大学举办。本课题负责人通过远程视频会议参加了此次交流会, 并就课题目前研究进展和相关问题与国际同行进行了深入交流。

7. 阶段性成果:

- 1) 建立软件经验库原型系统;
- 2) 完成国际会议论文 7 篇;
- 3) 提交专利申请 1 份, 完成软件著作权 3 份;
- 4) 培养硕士生 3 名;

下一年度工作计划, 包括国内外合作与交流计划

2011 年度的主要研究任务计划围绕以下三方面的内容展开:

- 探索 Semantic Web 技术驱动的经验软件工程方法:
经验软件工程的基本任务是通过分析来自软件工程界的数据来解决所研究的问题。而对于涉及其中的研究人员来说, 其基本任务是要理解过程数据, 制品及其相关关系的本质。针对软件开发经验性知识的重要性和稀缺性, 基于语义理解技术研究软件过程数据的自动收集和提取方法, 支持跨组织数据的共享、理解和重用, 同时进一步开展与 RPI 的合作;
- 设计软件过程演化本体框架:
结合前期研究基础和相关经验软件工程方法的资源, 设计和构造软件过程演化的基本概念模型、关系模型和演化模型, 构建软件过程演化本体框架;
- 构建基于本体的软件过程演化经验库
软件数据和经验对于经验研究不可或缺。结合研究的本体框架和数据提取方法, 基于已有数据资源, 研究构建可支持自动演化分析的软件过程演化经验库;

国内外合作与交流计划:

2011 年 5 月, 参加 ICSE 和 ICSP 国际会议;
2011 年 6-7 月, 邀请 RPI 丁力副教授来访;
2011 年 9 月, 参加 ISERN 和 ESEM 会议;
2011 年 10 月, 参加 COCOMO/软件成本建模论坛;

2011 年度预期成果:

在国际软件工程顶级期刊或会议上发表研究论文 1-2 篇。