

# 上海交通大学

## 2015 级硕士学位论文开题报告登记表

学号	115037910019	姓 名	俞飞樾	导 师	钟浩，沈备军
学科	软件工程	学院(系、所)	软件学院		
学位论文题目	并发程序修改模式的研究				
研究课题来源					
课题的意义以及研究的主要内容	<p>并发程序在现代软件开发活动中非常常见，也是学术界研究的热点问题之一。随着多核系统的不断普及，使用并发程序可以更好的利用多核系统的计算能力。然而，众所周知的是编写正确的并发程序是非常困难的，当并发地读写数据时，人们需要处理不同线程之间的复杂的同步问题。更糟糕的是，调试并发程序比调试一般的程序更具有挑战性，并发程序很难重现错误，每一次程序运行，线程调度通常是不一样的。</p> <p>本课题将从软件演化的角度来研究并发程序中的修改模式。近年来 <b>Github</b> 等开源代码托管平台逐渐流行，开源软件项目数呈爆发性增长趋势，在这些开源项目及其版本历史中蕴含着很多对今后软件开发有指导意义的信息。本课题将重点研究以下问题：</p> <ol style="list-style-type: none"> <li>1. 在开源的 <b>Java</b> 项目中，并发相关代码出现的频率如何？</li> <li>2. 对于一些常见的并发相关类库的使用，有什么趋势？</li> <li>3. 在并发程序中，有那些修改模式？</li> </ol> <p>修改模式指针对相似上下文的相似修改的提取与抽象</p> <ol style="list-style-type: none"> <li>4. 这些修改模式对并发程序软件开发与维护有什么帮助？</li> </ol> <p>大部分的并发程序研究集中在静态与动态分析，而对于历史信息的挖掘却没有那么深入。通过的并发程序演化历史的分析与挖掘，我们期望得到是否这些信息能够有效地帮助开发者更好的开发和维护并发程序。</p> <p>本课题的主要意义有：</p> <ol style="list-style-type: none"> <li>1. 从软件版本历史中，通过挖掘和归纳并发程序中的修改模式，对开发者开发和维护并发程序提供帮助。</li> <li>2. 对于具有代表性的并发程序的修改模式，在开源代码中寻找可以应用的场景并向开源社区提交补丁，期望能被接受，从而验证这些修改模式确实存在软件开发中存在，并能帮助今后的开发活动。</li> <li>3. 对于并发类库的设计者，可以从我的研究中的得到用户在使用一些并发特性时的倾向性，从而改进并发类库的设计。</li> <li>4. 研究表明并发程序的自动化工具支持依然不足，尽管 <b>findbugs</b> 等工具可以找到一些并发程序的隐患，但大部分的修改依然需要手动完成，我们期望具有更高自动化水平的工具，能够自动的探测出修改的模式并应用在合适的场景。</li> <li>5. 其他研究者可以考虑从历史代码中蕴含的信息的角度研究并发程序中的一些研究问题。</li> </ol>				



学 科 审 批 意 见	开题报告时间： 参加人数：教师 人，研究生 人。 审查结果： <input type="checkbox"/> 同意 <input type="checkbox"/> 不同意  <div style="text-align: right;">           签名：_____ 年 月 日         </div>
备 注	

## 上海交通大学软件学院硕士论文开题综述

### 1. 并发程序

Rui Gu(What Change History Tells Us about Thread Synchronization)等人从代码变化历史中得到有关线程同步的一些发现。他们研究了4个有代表性的开源C/C++项目，检查了超过250000个版本的代码，探究了有多少不同种类的代码修改，为什么做这些修改，何时做这些修改。他们发现了一些有趣的结论：(1)对于数量很可观的一部分关键区域（被锁保护的代码区域），大约20%到25%，锁是后来才加上的，这很不合理，通常情况下，代码和同步应该同时写好，说明程序员忘记了同步。当然，也有一些情况是由于程序的变化导致数据共享的方式发生了变化，需要加上同步。(2)在MySQL和Apache项目中，超过四分之三的关键区域发生了修改或移除，研究中超过四分之一的代码发生了同步方面的修改，例如关键区域别界的调整，锁变量的改变。可以看出，关于同步的修改很普遍，所以程序员需要一些辅助工具来帮助完成这些修改，来保证正确性，提高修改速度。(3)随着软件的发展演化，关键区域的修改数保持稳定，但这一修改数随着关键区域本身的发展演化而显著下降。和其他修改类型相比，为了提升性能或修复错误的修改通常发生在关键区域版本历史的早期。(4)修复正确性错误是关键区域修改的最主要的原因，为了提高性能而做的修改并没有那么多，但也不能忽略，大约10%的修改是由于这一原因。他们还研究了过同步问题和导致并发错误的原因。

### 2. 程序变换

上海交通大学软件学院  
20XX 年月日