**软件工程中心硕士研究生写作能力训练考核表**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **姓 名** | 段世凯 | | | **性 别** | 男 | | **学科专业** | 计算机软件与理论 |
| **入学年月** | 2014.09 | | **硕士指导教师** | | | 王伟 副研究员 | | |
| **类 别** | | 🞎 发表/录用论文 √ 技术报告 🞎 设计报告 | | | | | | |
| **题 目** | | GraphFlow:基于状态更新的流式图计算模型的设计与实现 | | | | | | |
| 论文/报告摘要：  图数据结构能够很好的表达数据之间的关联性，因此在社交分析、商品推荐、舆论监测和欺诈检测等应用中被广泛使用。而目前分布式处理框架Google Pregel、Spark GraphX和Flink Gelly等都是处理静态稳定的图数据；而已经提出的增量图计算模型和基于推测机制的并发更新模型，又分别受限于其串行的更新方式和节点更新只与节点接收消息相关的强力约束，使得模型在并发性低，表达能力有限。本文提出了一种基于状态更新的流式图计算模型GraphFlow，它将图数据的变化抽象成一系列的事件流，将用户关心的图计算结果抽象成图的状态，用户只需要定义图数据如何根据到达的事件增量式的进行状态的转换，就能够完成事件流到状态流的映射，提供实时反馈计算结果能力。通过对独立状态的分布式存储和并发更新策略，以及对关联状态的分区并行更新策略以及细粒度锁的更新策略，能够有效解决关联状态下更新冲突的问题，从而提高了系统的并行性。试验结果表明，GraphFlow系统的90%请求都能够在12ms内完成计算，且计算偏差在1%以内。  申请人签字：  年 月 日 | | | | | | | | |
| 审核小组总体评价：  是否通过： □ 优秀 □ 良好 □ 通过 □ 不通过  审核小组成员签名：  年 月 日 | | | | | | | | |