



## 软件研究所中长期发展规划

### 2009 年科技新星专项计划入选者

	姓 名:	李文波
	工作部门:	基础软件国家工程研究中心
	资助类别:	青年博士创新基金
	资助编号:	ISCAS2009-DR12
	资助金额:	12 万元
	支持周期:	2009 年 9 月-2012 年 9 月
研究方向	文本数据挖掘	
研究内容	<p>探索和挖掘概率主题模型的可并行性,系统性地研究和建立概率主题模型的可扩展求解算法,并将其应用到文本挖掘、信息检索等大规模文本处理任务。</p> <p>本项目具体的研究内容主要包括:(1)研究针对数据规模的具有可扩展性的高性能算法,关键问题是针对数据集分配的负载均衡算法;(2)研究针对模型规模的具有可扩展性的高性能算法,关键问题是模型分解算法;(3)研究基于众核平台的细粒度并行算法,关键问题是计算任务向众核上的硬件基础(SIMD)和计算模型(Stream)有效映射的算法;(4)大规模概率主题模型在信息检索和文本分类等领域中的应用。</p>	
预期成果	<p>系统研发:实现基于国产超级计算平台的高可用、高性能大规模概率主题模型求解系统;知识产权:提交研究相关的发明专利申请 1~2 项,或软件著作权申请 1~2 项;论文发表:在国内外重要学术期刊和学术会议上发表 3~5 篇研究相关的高质量论文;学术交流:参加研究相关的高水平国际国内会议 1~2 次,或参加研究相关的国际国内评测活动 1~2 次。</p>	