

自然科学奖公示信息模板

一、项目名称

多因素耦合复杂图模式挖掘模型理论与算法

二、提名者及提名意见

提名单位：陕西省教育厅

提名意见：该成果材料齐全、规范，无知识产权纠纷，人员排序无争议，符合陕西省科学技术进步奖提名条件。推荐该项目申报陕西省科学技术奖二等奖。

三、项目简介

复杂性科学被称为二十一世纪科学，已渗透到人类生产与生活的各个方面，包括金融经济、生态环境、医学治疗等。众多复杂系统产生了海量复杂网络数据，为分析与理解复杂系统的原理与机制带来了巨大的契机。因此，复杂网络数据分析和挖掘及其应用研究已成为国内外研究的焦点与热点，而备受关注。但是传统挖掘方法只关注复杂网络拓扑结构单一因素，忽视了时间、尺度、噪声、异质等多重因素对复杂系统的影响，难以有效反映与解释复杂系统。而多因素耦合复杂图模式可更加全面、有效地刻画与描述复杂系统，是理解与揭示复杂系统原理与机制的前提。

自 2010 年以来，课题组针对多因素耦合复杂图模式刻画难、挖掘难、应用难的问题进行系统性研究，构建了多因素耦合复杂图模式模型、研发了多因素耦合图模式挖掘算法、拓展了复杂耦合图模式的医学应用。在理论、算法与应用三个层面取得重要的基础性研究成果，主要创新包括：①多因素耦合复杂图模式模型与理论。针对时序网络动态演化图模式刻画难、理论缺失的问题，提出了面向时序网络的模块密度模型，量化了时序演化图聚类结构，证明了时序网络图聚类算法在目标函数上的等价性。针对复杂网络聚类量化与挖掘算法缺乏统一框架，提出了加权模块密度模型，证明了优化模块密度与聚类算法的等价性，将典型聚类簇结构量化和算法归约到统一框架下。②多因素耦合复杂图模式挖掘算法。针对在线动态环境中流数据模式动态性刻画与捕捉难的问题，提出了一种自动在线数据流分析算法，以自适应方式动态学习参数，刻画与捕捉动态环境中演化模式。针对时序网络拓扑结构不完整性，导致图模式挖掘难的问题，提出了一种图正则化非负矩阵分解算法，避免了时序网络压缩所导致的信息损失。针对现有多层网络聚类算法不能量化聚类簇层特异性，导致无法平衡聚类簇层内连通性与层间耦合性的问题，将多层网络聚类挖掘问题转化为联合优化问题，提出了面向多层网络聚类的线性判别学习算法，实现了多层网络聚类簇层特异性的量化。③多因素耦合复杂图模式医学应用。针对癌症恶化基因动态演化模式挖掘难与背景可解释性难的问题，提出了一种面向癌症恶化时序属性基因网络的非负矩阵分解算法，将基因时序网络和基因功能属性集成问题转化为多层属性网络聚类问题，实现了癌症恶化过程基因代谢路径的动态追踪。并应用于癌症甲基化基因多组学数据集成分析与新冠肺炎动态演化追踪，相关成果提交科技部并转交国家科技攻关小组疫情防控专班作为决策参考。研究成果为医学应用提供了有效的计算方法与分析手段。

四、客观评价

在国内外得到广泛关注，并且得到积极评价，包括美国科学院院士、欧洲科学院院士、ACM Fellow、IEEE Fellow 等国际与国内著名专家的广泛认可与积极评价。

五、代表性论文专著目录

序号	论文专著名称	刊名	作者	年卷 页码	发表 时间	通讯 作者	第一 作者
1	Evolutionary nonnegative matrix factorization algorithms for community detection	IEEE Transactions on Knowledge and Data	Xiaoke Ma, Di Dong	29(5): 1045-1058	2017/3/30	Xiaoke Ma	Xiaoke Ma
2	3E-LDA: three enhancements to linear discriminant analysis.	ACM Transactions on Knowledge Discovery from Data.	Yanni Li, BingLiu, Yongbo Yu, Hui Li, Jiaca Sun, Jiangtao Cui.	15(4): 57	2021/3/1	Hui Li	Yanni Li
3	ESA-Stream: efficient self-adaptive online data stream clustering	IEEE Transactions on Knowledge and Data	Yanni Li, Hui Li, Zhi Wang, BingLiu, Jiangtao Cui, Fei Hang	34(2): 617-630	2022/1/11	BingLiu, Jiangtao Cui	Yanni Li
4	Graph regularized nonnegative matrix factorization for temporal link prediction in dynamic networks.	Phyisca A	Xiaoke Ma, Penggang Sun, Yu Wang	496:121-136	2018/1/1	Xiaoke Ma, Penggang Sun	Xiaoke Ma
5	COVID-19 Evolves in Human Hosts	ACM SIGKDD '20 (CCF A 类会议)	Yanni Li, Bing Liu, ZhiWang, Jiangtao Cui, Kaicheng Yao, Pengfan Lv, Yulong Shen, Yueshen Xu, Yuanfang Guan, Xiaoke	https://doi.org/10.1145/1122445.1122456	2020/8/23	Xiaoke Ma	Yanni Li, Bing Liu

6	图数据挖掘理论、算法及应用	专著	马小科 刘晓刚	ISBN: 978-7- 5612- 8626-	2023/ 1/1	刘晓刚	马小科
---	---------------	----	------------	-----------------------------------	--------------	-----	-----

六、主要完成人情况

马小科，西安电子科技大学教授，项目的第一完成人，负责项目整体研究方案的创新构想与实施，复杂网络图模式挖掘模型、算法及其应用；是代表作1、4的第一作者兼(共同)通讯作者、代表作5的通讯作者、代表作6的第一作者。

董迪，中国科学院自动化研究所研究员，项目第二完成人。本项目中负责多因素耦合复杂图模式模型研究。是代表作1的参与人。

李雁妮，西安电子科技大学教授，项目第三完成人。本项目中负责多因素耦合复杂图模式模型与算法研究，是代表作2、3、5的第一作者。

刘晓刚，西北工业大学副教授，项目第四完成人。本项目中负责复杂图模式在医学的应用，是代表作6的通讯作者。

王玑，西安电子科技大学副教授，项目第五完成人。本项目中负责多因素耦合复杂图模式模型与算法研究，是代表作4的参与作者。

七、主要完成单位情况

西安电子科技大学：西安电子科技大学本项目申报单位，依托四项国家自然科学基金与一项陕西省自然科学基金项目进行科学攻关。西安电子科技大学是第二、三、四项国家自然科学基金项目，以及第一项陕西省自然科学基金项目的主持单位。项目第一、第三与第五完成人人依托西安电子科技大学全职科研，成果属于西安电子科技大学，其中西安电子科技大学是第一、第二、第四、第五篇代表作的第一单位兼通讯单位，也是第三、第六篇代表作（专著）的第一作者单位。

中国科学院自动化研究所：中国科学院自动化研究所是本项目的参与申报单位，是第一项国家自然科学基金项目主持单位。项目第二完成人依托中国科学院自动化所全职科研，成果属于所在单位，是第一篇代表作的参与单位。

西北工业大学：西北工业大学是本项目的参与申报单位，给予了项目必要的经费、人员等支撑条件，为本项目提供了相关研究平台。项目第四申报人依托西北工业大学全职工作，成果属于所在单位，是第六篇代表作的通讯单位。

八、完成人合作关系说明

项目申报人马小科、董迪、李雁妮、刘晓刚、王玑，自2008年起进行深度科研联合攻关，具体合作关系如下：

（1）项目第一完成人马小科与第二完成人董迪自2015年起开始医工合作，合作发表代表作1等多篇论文。

（2）2015年项目第一完成人马小科全职回国工作，与项目第三完成人李雁妮隶属于同一课题组，合作发表代表作5等多篇论文。

（3）项目第一完成人马小科与项目第四完成人刘晓刚师从张远平教授攻读硕士学位，双方合作发表代表作6等多篇论文。

（4）项目第一完成人马小科与项目第五完成人王玑师从高琳教授攻读博士学位，双方合作发表代表作4等多篇论文。