

# 上海市自然科学奖公示材料

## 项目名称：动力学驱动的数据科学理论和方法研究

### 简介：

基于海量时序列数据，我们研究了高维复杂动力系统在临界突变前的数据特征与普适的动态行为，发现了变量的时空关联与调控关系，提出了在高维、大噪声和个体化小样本等严苛数据条件下的动态网络标志物理论，开发了行之有效的计算手段。在数据降维与降噪、分子网络推断、特别是在复杂系统临界突变预警和控制方法等方面取得了突破性成果。研究中我们重视数据的可视化、理论的总结和算法开发，催生了一系列高效的动态特征挖掘与分析方法，关键模块建模与系统重构方法，决策制定与动态优化方法。这些成果已经被国内外团队广泛应用于癌症转移与复发，公共卫生实时监测、亚健康状态预警、时间序列预测等方面，对动力学驱动的数据科学理论和算法发展起到了重要的推动作用。本项目主要取得了以下成果：

(1) 复杂疾病过程的临界理论及动态网络标志物方法：在国际首次建立了复杂疾病及动态生物过程的临界状态的预测方法和理论，特别是建立了基于大数据检测临界状态的必要条件及其动态网络标志物(DNB: dynamic network biomarker)，为该领域开辟新方向。该成果不仅已成功应用到肝癌转移前兆诊断、糖尿病临界检测及药物拮抗动态过程等研究，而且也应用到生态系统及金融系统等的风险分析和临界预测。

(2) 提出了全新的概念，即“偏独立性”(PMI: part mutual information)以及新的度量方法：PMI不仅能够克服条件互信息的缺点，同时还能保持着互信息和条件互信息的性质。特别是首先定义了部分互信息用于量化变量之间直接关联性，然后从理论上给出部分互信息和互信息以及条件互信息之间的内在联系该工作对于大数据网络研究提供了全新的概念和理论，不仅可用于生物分子网络的高精度构建，而且可以应用于其他网络(领域)中各种复杂因素的因果关系推断。

(3) 小样本高维度数据的生物分子网络构建理论和方法：提出了基于转录组学数据的基因调控网络推导数学模型和方法，有效缓解了小样本高维度(小  $n$ ，大  $p$ )的困难，解决了直接和间接调控关系难以区分的问题，提出了有效抽取序列和结构等多尺度数据中分子相互作用模式的特征抽取和选择算法，显著降低了分子相互作用预测及网络构建中的假阳率,为基于数据的生物分子网络构建提供全新的方法。

(4) 首次研究了非线性复杂网络控制的时间及能量问题，设计了新的闭环控制器，利用严格的理论分析给出了网络控制时间及能量成本上界的解析估计。该论文提出的方法具有广泛的应用范围，可以用来研究生态网络、细胞网络以及神经网络的控制问题。

五篇代表论文发表在 *Proceedings of the National Academy of Sciences*, *Physical Review Letters* 和 *Bioinformatics* 等国际著名学术期刊，被 SCI 他引 629 次，入选 ESI 高被引论文 2 篇。完成和承担国家重点研发计划 1 项(首席)、国家自然科学基金重点项目 5 项(主持)、面上项目 7 项(主持)。完成人多次应邀在国际学术会议上做大会或主旨报告，担任“香山”科学会议和“冷泉港”科学会议在内的多个国内外学术会议的主席，成员中获得国家基金委“杰青”一人次，国家基金委“优青”两人次。

**主要完成单位：**中国科学院分子细胞科学卓越创新中心；复旦大学；中国科学院数学与系统科学研究院；华南理工大学；

**主要完成人：**陈洛南、刘锐、林伟、王勇

**提名等级：**一等奖

**提名单位：**中国科学院上海分院

## 代表性论文

1. Juan Zhao, Yiwei Zhou, Xiujun Zhang, Luonan Chen. Part mutual information for quantifying direct associations in networks. *Proceedings of the National Academy of Sciences*, 2016, 113(18): 5130-5135.  
SCI 他引次数：70
2. Rui Liu, Xiangdong Wang, Kazuyuki Aihara, Luonan Chen. Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. *Medicinal Research Reviews*, 2014, 34(3): 455-478.  
SCI 他引次数：106      ESI 高被引论文
3. Yongzheng Sun, Siyang Leng, Ying-Cheng Lai, Celso Grebogi, Wei Lin. Closed-loop control of complex networks: A trade-off between time and energy, *Physical Review Letters*, 2017, 119, 198301  
SCI 他引次数：15
4. Luonan Chen, Rui Liu, Zhi-Ping Liu, Meiyi Li, Kazuyuki Aihara. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network

biomarkers. *Scientific Reports*, 2012, 2:342.

SCI 他引次数: 218      [ESI 高被引论文](#)

5. Yong Wang, Trupti Joshi, Xiang-Sun Zhang, Dong Xu, Luonan Chen, Inferring gene regulatory networks from multiple microarray datasets. *Bioinformatics* 22 (19), 2413-2420

SCI 他引次数: 220      [ESI 高被引论文](#)