

基于 LDA 主题模型的文献关联分析及可视化研究

王 丽 邹丽雪 刘细文

(中国科学院文献情报中心 北京 100190)

(中国科学院大学 北京 100049)

摘要:【目的】基于机器学习进行有效的科研主题发现与分析建模研究。【方法】以 LDA 主题模型为基础进行科研主题发现,通过 Python 语言实现分析建模,关联分析科研主题的年代、机构、基础研究与应用研发等信息,对比披露了各年代、机构的研发重点及基础研究和应用研究的差异,并对结果进行可视化。【结果】基于101 813 条石墨烯领域的论文和专利数据进行实证分析,结果表明通过建模进行12个主题的发现、关联分析及其可视化全流程可在很短时间(约2分钟)内完成,通过机器学习的手段突破了传统分析工具以及数据分析体量的局限性,提升了针对大量文本数据信息挖掘及分析的效率。【局限】尚未探讨主题间关联关系权重及主题演进等涉及复杂网络分析的内容。【结论】机器学习对情报分析有着无限可能,需要基于实际需求开发更多有效的分析模型,满足大文本分析和可视化。

关键词: LDA 主题模型 数据分析 机器学习 Python 数据可视化

分类号: TP393

DOI: 10.11925/infotech.2096-3467.2017.1058

1 引言

数字科研、数字教育和数字信息网络正在全面改变着传统信息服务模式,文献情报服务正处于颠覆性变革阶段,大数据情报分析和知识服务时代已经来临。基于数据科学和计算科学的数据处理和分析方法为情报分析及信息服务带来新的契机。

以往的情报分析大多基于现成工具进行数据处理、分析以及可视化,导致情报分析受限于工具的预设功能和数据处理能力,对于分析人员的信息挖掘和表达形成了客观的局限性。近年来随着科研成果数字化,论文和专利数据作为文献情报分析的主体急剧增加,现有工具的处理能力已经无法承载产业级数据的分析,急需探索针对大文本数据分析的新方法和新途径。深度信息发现是通过挖掘和比较不同层次的关联关系实现的,现有分析工具可以满足常规的分析需求,

然而自定义和多维度的分析往往受限于其通用性。针对上述问题,本文以石墨烯产业为例探索大数据文本分析方法,以 LDA (Latent Dirichlet Allocation)主题模型为基础挖掘石墨烯产业研发热点方向,并建立机构、年代、类型等关联分析及可视化模型。

机器学习中的无监督学习主要依赖于统计模型,主题模型则属于无监督学习。LDA 主题模型是一种无监督机器学习算法,以词、主题和文档三层贝叶斯概率为核心结构发现文本隐含模式。LDA 通过使用联合概率分布计算给定观测变量值下的隐含变量的条件分布,核心部分如公式(1)^[1]所示。

$$p(w|\theta) = \sum_{z=1}^k p(w|z)p(z|\theta) \quad (1)$$

其中, w 表示词, $|V|$ 表示所有词的个数; z 表示主题, k 表示隐含主题个数; $W = \langle w_1, \dots, w_N \rangle$ 表示文档, 一个文档有 N 个词; θ 是主题向量, 向量的每一列表

通讯作者: 王丽, ORCID: 0000-0002-9513-6159, E-mail: wangli@mail.las.ac.cn。

示每个主题在文档出现的概率, 即 θ_i 揭示文档属于主题 i 的可能性, 该向量为非负归一化向量 ($\theta_i \geq 0, \sum_i \theta_i = 1$), $p(\theta)$ 服从 Dirichlet 分布; $p(w|z)$ 表示给定 z 时 w 的分布, 是一个 $k \times |V|$ 矩阵, 每行表示这个主题对应单词的概率分布; $p(z|\theta)$ 表示给定 θ 时主题 z 的概率分布; $p(w|\theta)$ 表示给定 θ 时词 w 的概率分布。从文档推断隐含主题结构是一种计算后验分布的计算, 公式(1)中, w 为给定观测变量, θ 和 z 为隐藏变量, LDA 后验分布计算在给定观测文档下计算隐含变量的条件分布。

LDA 主题模型被广泛应用于文本挖掘领域, 如文本主题识别、文本分类以及文本相似度计算等方面。Lee 等^[2]基于专利数据在 IPC 共现分析基础上利用 LDA 模型发现新兴技术主题; 王博等^[3]基于专利数据利用 LDA 模型实现专利主题和机构之间的关联分析, 但未进行可视化研究; 任智军等^[4]基于 LDA 模型研究新兴技术识别方法, 对论文和专利数据进行整合, 基于 LDA 模型获取技术主题, 并未对可视化进行过多研究; 杨超等^[5]对基于 SAO 结构的 LDA 模型方法进行研究, 致力于构建 SAO 语义模型与 LDA 统计模型相结合的专利技术主题分析方法, 并未对可视化和关联分析有过多探讨; Suominen 等^[6]基于专利数据通过 LAD 模型对企业研发方向进行分析和预测, 并基于 Gephi 进行主题可视化, 未进行专利和论文整合数据的对比研究。

2 研究思路与框架

本文利用美国化学文摘社 (Chemical Abstracts Service, CAS) 提供的石墨烯的相关论文及专利文献, 截至 2016 年 4 月 5 日, 共获取到 78 756 篇论文, 23 057 项专利。原始文件为 XML 格式, 为了便于后续分析, 人工解析提取有效分析字段, 并将其转换为常用的 CSV 格式文件。开发数据解析工具将原始数据处理为结构化数据, 对作者、机构进行归一化处理, 保证后续分析结果的准确性。

论文数据和专利数据从结构、长度、语言的表达习惯都存在很大差异, 在进行混合数据分析时, 寻找等质等量语料结构用于主题模型建模是关键。CAS 对出版物共同标引, 如主题标引、概念标引、物质标引等, 为本研究进行专利和论文混合数据分析提供很大便利。选取 CAS 的主题标引将单篇文档转换为多个主题词所构成的集合, 用于主题模型建模分析。

采用 Gibbs 采样进行 LDA 建模, 形成两个矩阵, 即文档-主题概率矩阵和特征词-主题概率矩阵, 基于上述两个矩阵通过 Python 建立分析模型, 展开后续的关联分析, 如主题年代演进对比分析、机构的研究主题对比分析、主题研发对比分析等。

基于词-主题概率矩阵通过 Python 建模实现 LDA 主题聚类的可视化, 在进行关联分析的同时实现可视化表达。研究框架和分析流程如图 1 所示。

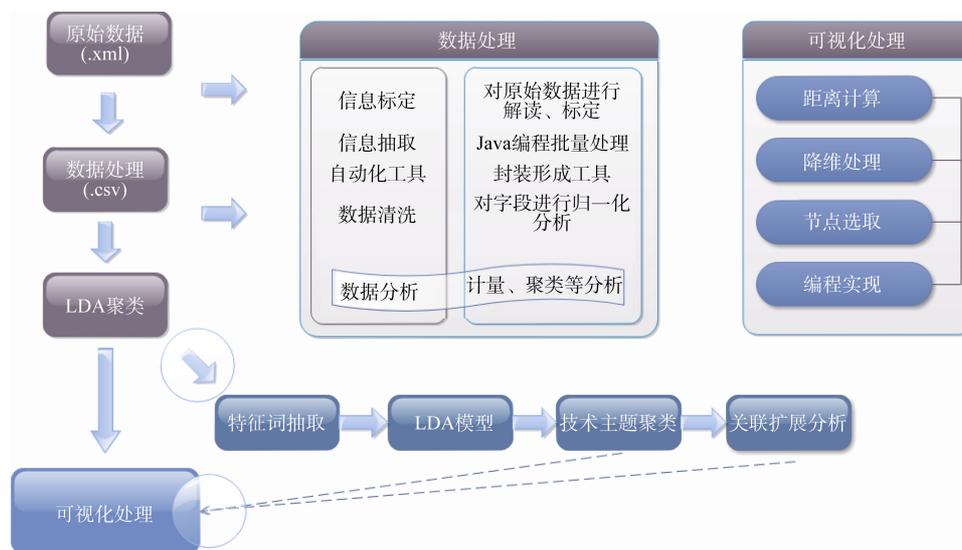


图 1 基于 LDA 模型的研究框架和流程(以 CAS 数据为例)

3 LDA 主题模型分析及可视化

对数据进行预处理, 整合论文和专利数据, 提取文献号、年代、机构、主题词、物质、功能等特征字段, 形成分析数据。基于 LDA 主题模型, 对 101 813 条专利和论文混合数据进行主题聚类, 由于后续涉及

聚类主题的人工判读, 实验阶段分析颗粒度为主题大类划分, 因此参数设置为 12-20 个主题, $\alpha=4.1$, $\beta=0.01$, 进行 1000 次迭代。经过对主题之间的交叉判读, 最终参数选定为 12 个主题, 每类主题 Top10 的特征词及其概率分布如表 1 所示。

表 1 石墨烯领域主题-特征词分布

主题 1 石墨烯建模仿真		主题 2 石墨烯电化学性能		主题 3 石墨烯 FET 器件		主题 4 石墨烯反常霍尔效应	
特征词	概率	特征词	概率	特征词	概率	特征词	概率
monolayers	0.040	surface structure	0.075	electric current-potential relationship	0.100	magnetic field effects	0.037
simulation and modeling	0.033	nanoparticles	0.073	electric conductivity	0.052	electric conductivity	0.030
multilayers	0.024	cyclic voltammetry	0.051	electric resistance	0.050	electron transport	0.027
electric field effects	0.022	nanocomposites	0.049	electric capacitance	0.043	nanoribbons	0.026
phonon	0.017	glassy carbon electrodes	0.033	electrodes	0.043	band structure	0.025
electric conductivity	0.016	Nano sheets	0.032	field effect transistors	0.038	fermi level	0.023
electric current carriers	0.015	electron transfer	0.022	double layer capacitors	0.035	quantum hall effect	0.020
optical transmission	0.014	electric impedance	0.021	raman spectra	0.034	landau level	0.018
semiconductor materials	0.014	x-ray diffraction	0.020	solar cells	0.023	magnetization	0.017
dielectric constant	0.013	ph	0.019	electric impedance	0.019	tight-binding method	0.015
主题 5 石墨烯/碳纳米管复合材料		主题 6 石墨烯生物相容性		主题 7 氧化石墨烯		主题 8 石墨烯高分子复合材料	
特征词	概率	特征词	概率	特征词	概率	特征词	概率
films	0.045	human	0.031	reduction	0.039	polyesters	0.035
chemical vapor deposition	0.036	electronic device fabrication	0.021	oxidation	0.036	carbon nanotubes	0.034
carbon nanotubes	0.034	surface treatment	0.016	adsorption	0.033	epoxy resins	0.022
annealing	0.026	homo sapiens	0.014	surface area	0.033	polysiloxanes	0.020
electric conductors	0.025	chemically modified electrodes	0.014	Nano sheets	0.017	polyimides	0.020
etching	0.020	ph	0.014	nanostructured materials	0.017	polyurethanes	0.019
metals	0.020	quantum dots	0.013	sonication	0.017	polyamides	0.018
coating process	0.020	stability	0.013	exfoliation	0.016	polyoxyalkylenes	0.015
electrodes	0.018	fluorescence	0.012	pore size distribution	0.016	polyethers	0.015
sheet resistance	0.018	nanoscale surface modification	0.012	pore size	0.015	coating materials	0.014

(续表)

主题 9 石墨烯纳米带		主题 10 石墨烯复合材料的光学性		主题 11 石墨烯复合材料的力学性		主题 12 石墨烯储能电池	
特征词	概率	特征词	概率	特征词	概率	特征词	概率
density of states	0.057	raman spectra	0.093	nanocomposites	0.049	secondary batteries	0.063
band gap	0.051	microstructure	0.044	thermal conductivity	0.031	carbon nanotubes	0.057
nanoribbons	0.044	x-ray photoelectron spectra	0.041	thermal stability	0.030	composites	0.053
density functional theory	0.043	nanoparticles	0.041	polymer morphology	0.028	fluoropolymers	0.047
band structure	0.036	nanocomposites	0.038	young's modulus	0.028	carbon black	0.035
electronic structure	0.035	surface structure	0.037	electric conductivity	0.024	battery anodes	0.034
binding energy	0.026	nanostructures	0.036	carbon nanotubes	0.020	lithium-ion secondary batteries	0.023
electron density	0.023	uv and visible spectra	0.032	tensile strength	0.020	heat treatment	0.020
fermi level	0.020	nanosheets	0.029	molecular dynamics simulation	0.020	battery cathodes	0.015
band structure	0.020	ir spectra	0.025	strain	0.017	carbon fibers	0.015

基于 LDA 主题模型分析得到 12 个主题, 通过分析出现频次较高的主题特征词对各主题进行解读标识。将每个文档根据文档-主题概率进行主题归类处理, 为后续建模关联分析奠定基础, 即时间-主题模型的主题年代演进对比分析、机构-主题模型的机构研究主题对比分析、类型-主题-时间模型的主题研发对比分析等。

根据 LDA 主题模型分析形成的特征词-主题概率, 进行特征词相似度计算, 由于以主题概率作为特征词向量, 本研究采用余弦相似度算法计算特征词的空间距离, 如公式(2)所示。

$$D = 1 - \text{cosine_similarity}(w_i, w_j) = 1 - \frac{w_i \cdot w_j}{\|w_i\| \|w_j\|} \quad (2)$$

其中, D 是特征词的空间距离, w_i, w_j 分别代表主题概率向量。

通过多维缩放(MultiDimensional Scaling, MDS)对空间距离进行降维处理的同时保留数据空间距离的差异性, 利用 Python 的 Matplotlib 绘图工具包建模实现基于 LDA 主题模型聚类的可视化。由于聚类可视化显示了各个点所代表的特征词, 为了视觉辨析性选择特征词概率 0.0035 作为显示阈值, 结果如图 2 所示。

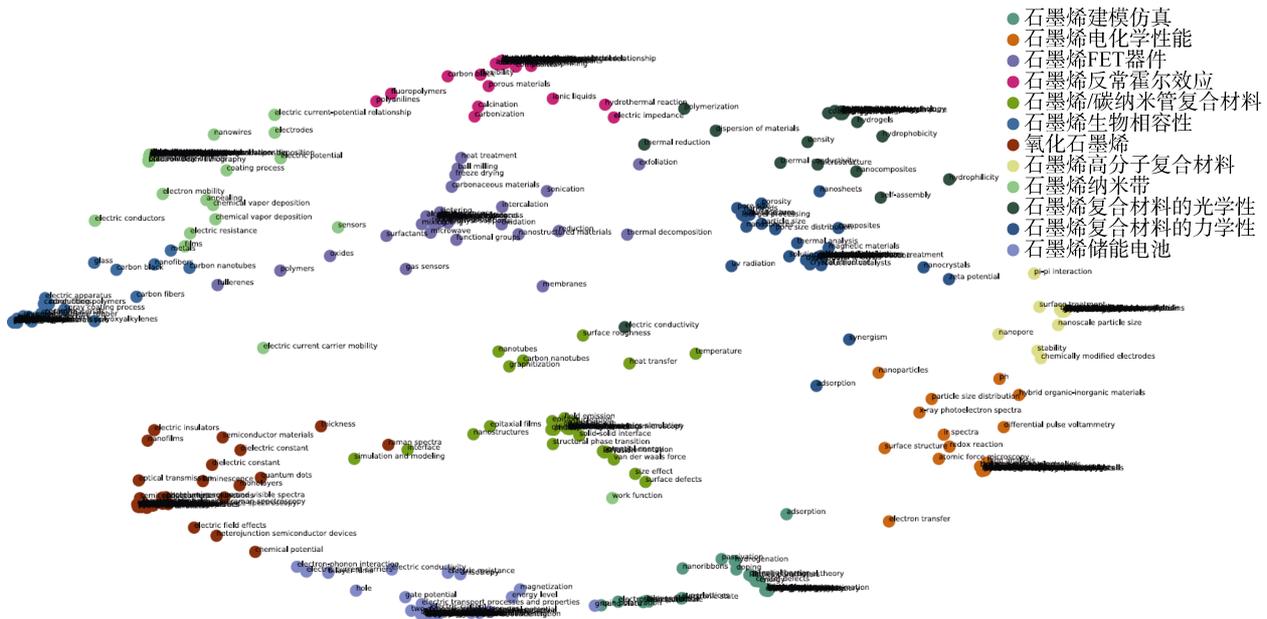


图 2 石墨烯领域主题分布

4 关联分析及可视化

基于 LDA 主题模型进行大体量文本数据的主题发现和关联分析以及结果可视化将涉及到大量的计算。本研究基于 Python 的 Numpy、Pandas 等工具包进行分析建模, 并利用 Matplotlib 绘图工具包建模实现分析结果的可视化, 解决绘图标签中文显示乱码的问题。通过反复优化和测试, 最终从 101 813 条数据的

读取、12 个隐含主题的发现及可视化到本研究涉及的关联分析及可视化, 整个流程所需时间约 2 分钟左右, 提高了执行速度且节省了内存空间。

4.1 时间-主题模型

基于 LDA 主题模型构建时间-主题模型, 通过分组循环计算各个主题领域的各年代文档数量, 实现主题年代演进对比分析及可视化, 如图 3 所示。

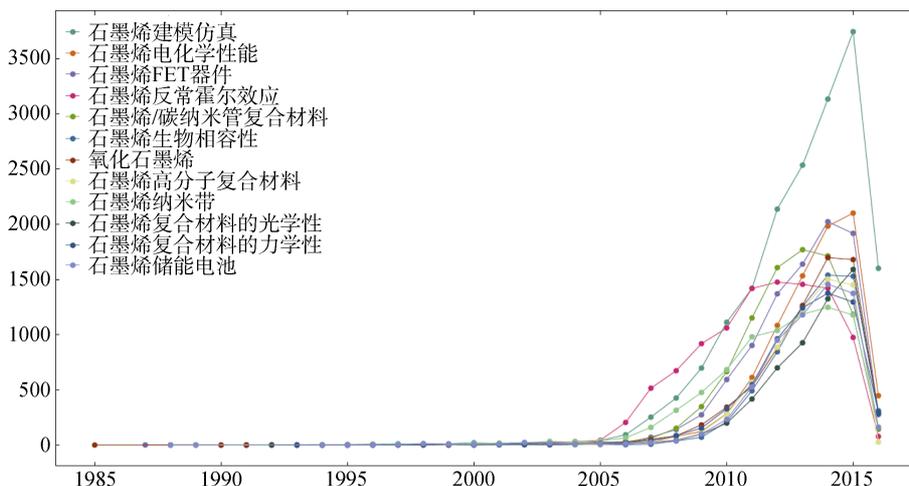


图3 石墨烯领域主题年代演进对比分析

根据本次聚类结果对石墨烯领域各主题年代演进对比分析发现: 第一个主题“石墨烯建模仿真”在近年的研发产出激增; 第二个主题“石墨烯电化学性能”的研究持续增长但增速减缓; 第三个主题“石墨烯 FET 器件”的研究在 2014 年-2015 年有所波动; 第四个主题“石墨烯反常霍尔效应”和第五个主题“石墨烯/碳纳米管复合材料”的研究产出下浮较大。

4.2 机构-主题模型

基于 LDA 主题模型构建机构-主题模型, 进行 12 个主题的 Top10 机构分析及可视化, 如图 4 所示。机构-主题模型建立的基本思路是通过分组循环计算各个主题领域的研究机构并进行降序处理和提取, 实现主题机构分析; 主要机构的研究主题对比分析则首先对石墨烯领域的主要机构进行提取, 然后循环计算各个主题领域的研究机构, 最后和主要机构进行融合计算, 从而实现主要机构的研究主题对比分析。

根据聚类结果对石墨烯领域各主题的 Top10 机构分析发现: 除第 4 主题“石墨烯反常霍尔效应”外, 中

国科学院在各主题的研究产出均位列第一。各主题的研究机构以科研单位和大学为主。此外, 第 5 主题“石墨烯/碳纳米管复合材料”和第 8 主题“石墨烯高分子复合材料”的 Top10 机构出现公司, 可见石墨烯/碳纳米管复合材料、石墨烯/高分子复合材料方面的研究受到企业的关注。当然, 由于论文数量远超专利数量, 在本次混合数据分析中以申请专利为主的公司其排名难免靠后。

根据聚类结果对石墨烯领域 Top10 机构的 12 个主题进行对比分析, 如图 5 所示。分析发现各机构研发各有偏重: 中国科学院在“石墨烯建模仿真”、“石墨烯/碳纳米管复合材料”研究产出突出, 在“石墨烯反常霍尔效应”及“石墨烯纳米带”方面的研究产出相对较少; 美国加州大学在“石墨烯建模仿真”、“石墨烯反常霍尔效应”及“石墨烯 FET 器件”方面的研究产出较多; 清华大学在“石墨烯 FET 器件”方面研究产出相对突出; 新加坡南洋理工大学在“石墨烯建模仿真”、“石墨烯电化学性能”及“石墨烯 FET 器件”方面研究产出较

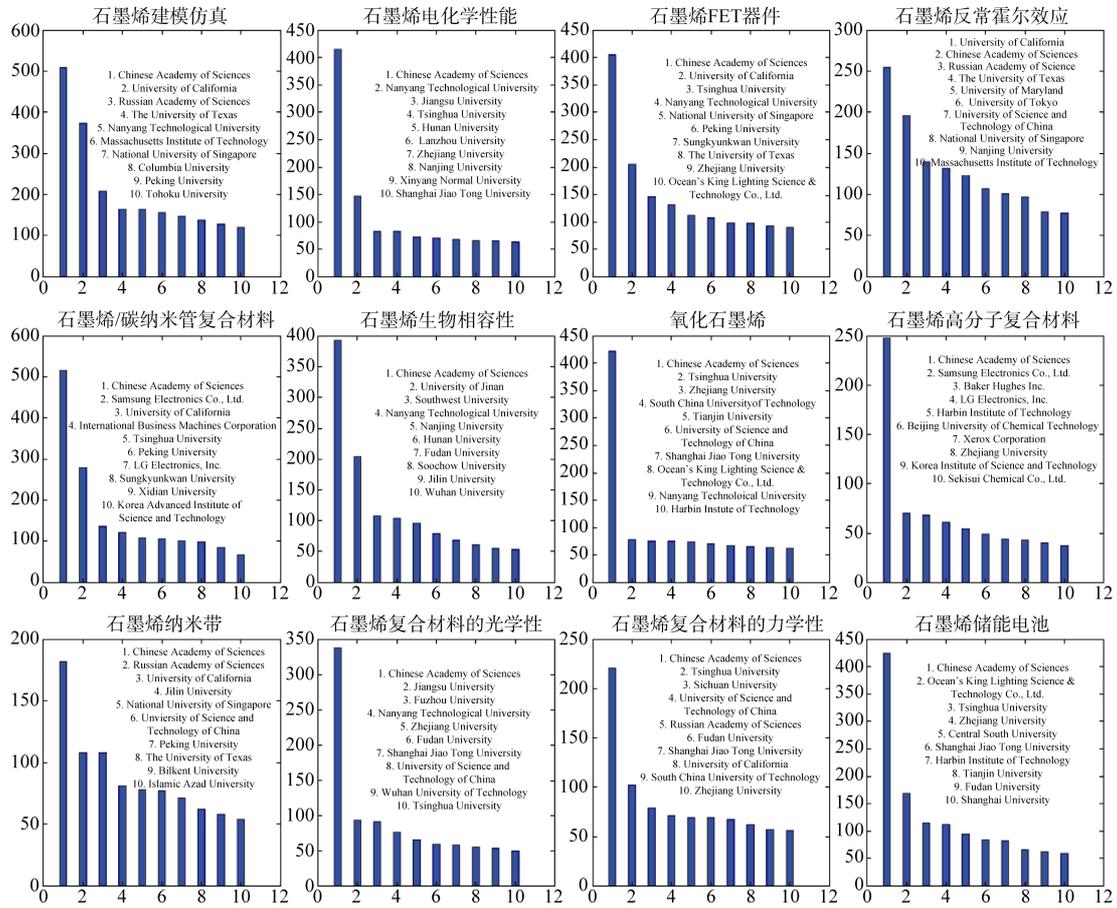


图 4 石墨烯领域 12 个主题的 Top10 机构分析

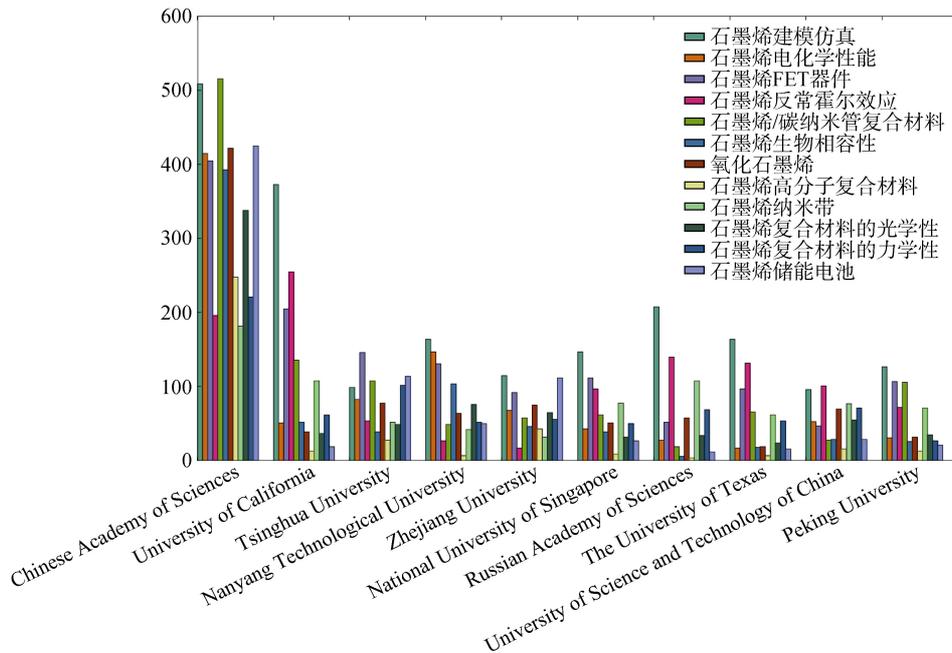


图 5 石墨烯领域 Top10 机构的 12 个主题对比分析

多; 浙江大学的产出主要集中在“石墨烯建模仿真”和“石墨烯储能电池”方面; 南京大学、俄罗斯科学院、美国德州大学、中国科技大学除了在“石墨烯建模仿真”方向产出较多, 皆在“石墨烯反常霍尔效应”方向研究产出较多; 北京大学在“石墨烯建模仿真”、“石墨烯 FET 器件”及“石墨烯/碳纳米管复合材料”方面研究产出相对较多。

4.3 类型-主题-时间模型

基于 LDA 主题模型构建类型-主题-时间模型, 进行主题基础研究和技术研发对比分析并可视化, 如图 6 所示。首先对文档进行专利和论文两个类别的分类处理, 然后循环计算每个主题下各年代的文档数量, 最后分类计算并融合各年代不同类别的文档数量。

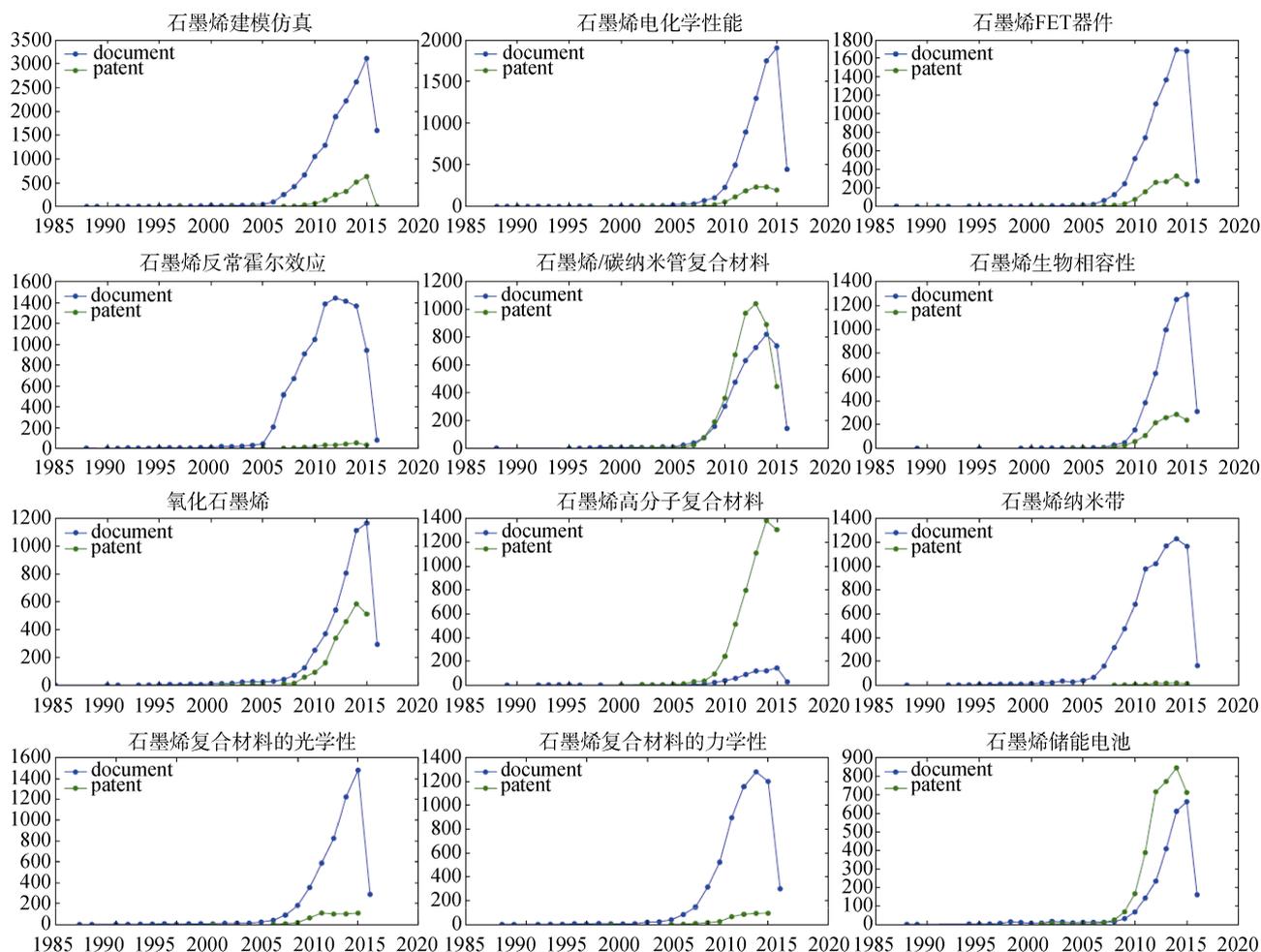


图 6 石墨烯领域研发对比分析

对石墨烯领域基础研究和应用研发进行对比分析发现: 第 5 主题“石墨烯/碳纳米管复合材料”和第 12 主题“石墨烯储能电池”的专利产出量居论文产出量之上但差距较小, 可见在这两个方面的相关研究已经从基础研究迈向应用研发, 同时基础研究持续跟进以待突破。第 8 主题“石墨烯高分子复合材料”的专利产出量高于论文产出, 与石墨烯作为添加剂广泛应用的现状相符。第 7 主题“氧化石墨烯”尽管

论文产出量高于专利, 专利产出仍然可观, 可见氧化石墨烯也有一定的应用市场, 但研究上也待进一步突破。第 4 主题“石墨烯反常霍尔效应”、第 9 主题“石墨烯纳米带”、第 10 主题“石墨烯复合材料的光学性”以及第 11 主题“石墨烯复合材料的力学性”等方面的研究目前还是以基础研究为主。第 3 主题“石墨烯 FET 器件”和第 6 主题“石墨烯生物相容性”等方面的研究以基础研究为主, 应用研发逐步跟进,

正如基于石墨烯的集成电路与半导体器件将给现有的半导体工业注入新的活力; 石墨烯生物传感器用于生物检测、生理监测、疾病诊断、环境及食品检测等领域的应用潜力逐步增强。

5 结 语

使用机器学习建模分析, 可以摆脱分析工具在处理体量及分析模型上的局限, 根据分析目的选取最优算法进行模型的改进和拓展, 同时可根据所要表达的信息进行可视化而不限制思维。采用 LDA 主题模型对专利和论文混合数据进行聚类分析及可视化, 基于此优化算法建立时间-主题、机构-主题、类型-主题-时间模型, 并进行关联分析及可视化。通过 CAS 提供的石墨烯相关论文及专利文献数据进行模型应用和检验, 由于 LDA 模型的概率采样每次重新生成主题的结果不尽相同, 但在对石墨烯 12 个主题的多次重复生成过程中发现, “石墨烯建模仿真”、“石墨烯 FET 器件”、“石墨烯反常霍尔效应”、“石墨烯生物相容性”、“氧化石墨烯”、“石墨烯高分子复合材料”、“石墨烯/碳纳米管复合材料”、“石墨烯储能电池”、“石墨烯复合材料力学性”等主题较为稳定。

尽管实证数据为 CAS 提供的数据, 但由于 LDA 模型是统计模型, 因此本文构建的分析模型可适用于其他类型数据。值得注意的是, 专利论文混合类型数据需要对分析语料进行预处理, 保证混合数据语料表达形式及数据宽度的一致性。如采用 Derwent 加工后专利摘要和任意数据库论文摘要的混合数据作为语料, 可直接采用本文的分析模型, 但是如果是专利原始文件和论文的混合数据, 由于专利与论文的表达形式差异需要进行语料预处理。未来将继续探索其他模型方法改进混合数据语料预处理模块, 同时针对分析结果探索最优的可视化模型也是未来的研究方向。

参考文献:

- [1] Blei M D, Ng Y A, Jordan I M. Latent Dirichlet Allocation[J]. Journal of Machine Learning Research, 2003, 3: 993-1022.
- [2] Lee W S, Han E J, Sohn S Y. Predicting the Pattern of Technology Convergence Using Big-Data Technology on Large-Scale Triadic Patents[J]. Technological Forecasting & Social Change, 2015, 100: 317-329.
- [3] 王博, 刘盛博, 丁堃, 等. 基于 LDA 主题模型的专利内容分析方法[J]. 科研管理, 2015, 36(3):111-117. (Wang Bo, Liu Shengbo, Ding Kun, et al. Patent Content Analysis Method Based on LDA Topic Model[J]. Science Research Management, 2015, 36(3): 111-117.)
- [4] 任智军, 乔晓东, 张江涛. 新兴技术发现模型研究[J]. 现代图书情报技术, 2016(8): 60-69. (Ren Zhijun, Qiao Xiaodong, Zhang Jiangtao. Discover Emerging Technologies with LDA Model[J]. New Technology of Library and Information Service, 2016(8): 60-69.)
- [5] 杨超, 朱东华, 汪雪锋, 等. 专利技术主题分析: 基于 SAO 结构的 LDA 主题模型方法[J]. 图书情报工作, 2017, 61(3):86-96. (Yang Chao, Zhu Donghua, Wang Xuefeng, et al. Technical Topic Analysis in Patents: SAO-based LDA Modeling [J]. Library and Information Service, 2017, 61(3): 86-96.)
- [6] Suominen A, Toivanen H, Seppänen M. Firms' Knowledge Profiles: Mapping Patent Data with Unsupervised Learning [J]. Technological Forecasting & Social Change, 2017, 115: 131-142.

作者贡献声明:

王丽: 提出研究思路, 设计研究方案, 模型建立, 代码实现, 撰写及修改论文;
邹丽雪: 数据清洗, 论文修改;
刘细文: 提出研究思路。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: wangli@mail.las.ac.cn。
[1] 王丽, 邹丽雪. graphene.xml. 美国化学文摘提供的石墨烯原始数据。
[2] 王丽, 邹丽雪. graphene.csv. 提炼后的石墨烯分析数据。
[3] 王丽. graphene.ipynb. 石墨烯分析模型代码。

收稿日期: 2017-10-24
收修改稿日期: 2017-12-09

Visualizing Document Correlation Based on LDA Model

Wang Li Zou Lixue Liu Xiwen

(National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

(University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: [Objective] This paper tries to construct data analysis model for the topics of scientific research based on machine learning. [Methods] First, we clustered data with the Latent Dirichlet Allocation model. Then, we investigated the correlation among year, institution and research types with the help of Python modules. Finally, we revealed and visualized the key research areas of every year or institution. [Results] We analyzed 101,813 papers and patents of graphene industry research. The proposed method finished the topic identification, correlation analysis, and visualization in about two minutes. [Limitations] More research is needed to explore the network analysis issues. [Conclusions] Machine learning provides enormous potentiality for intelligence studies, especially the large volume text analytics and visualization.

Keywords: LDA Model Data Analysis Machine Learning Python Data Visualization

机器学习服务比较: 谷歌、微软、亚马逊

服务型导向的机器学习(Machine Learning as a Service, MLaaS) 是自动化和半自动化云平台的一个概括性定义, 这个云平台范围很广, 包含数据预处理、模型训练、模型评估、以及预测。而且预测结果可以通过 REST API 与其他 IT 系统进行交互。

亚马逊机器学习服务、微软 Azure 机器学习和谷歌云 AI 是三个领先的云 MLaaS 服务, 只有少量甚至没有数据科学专业知识的用户也能利用这些服务进行快速模型训练和部署。

在用于自定义预测分析任务的机器学习服务方面, 微软 Azure 机器学习工作室拥有最全面的服务; 然而, 如果需要用到深度神经网络, 亚马逊的 SageMaker 和谷歌的机器学习引擎则更加合适。亚马逊机器学习服务有两个层面: 亚马逊机器学习的预测性分析和数据科学家的 SageMaker 工具。微软 Azure 提供的服务可以分为两个主要类型: Azure 机器学习工作室和机器人服务。谷歌的机器学习引擎适合有经验的数据科学家, 它非常灵活, 并建议将 TensorFlow 的云端基础架构作为机器学习的驱动程序。

除了成熟的平台, 用户还能使用机器学习 API, 这些 API 提供了底层训练好的模型服务, 用户可以输入数据, 然后获取结果。API 完全不需要机器学习的专业知识。目前, 这三家供应商的 API 大致可以分为三大类: 文本识别、翻译和文本分析; 图像+视频识别和相关分析; 其他, 包括特定的未分类服务。

亚马逊针对语音和文本处理而提供的 API 包括: 亚马逊 Lex、亚马逊 Transcribe、亚马逊 Polly、亚马逊 Comprehend, 以及亚马逊 Translate。针对图像和视频处理提供的 API 是亚马逊 Rekognition。微软针对语音、文字处理, 图像与视频处理提供的 API 是 Azure 认知服务。谷歌针对语音、文字处理, 图像与视频处理提供的 API 是谷歌云服务, 包括自然语言云处理接口、云语音接口、云翻译接口、云视觉接口、云视频智能。

(编译自: <https://www.altexsoft.com/blog/datascience/comparing-machine-learning-as-a-service-amazon-microsoft-azure-google-cloud-ai/>)

(本刊讯)