相关文档探测方法在科技查新中的应用研究

曹 燕 何晓敏 陈 亮 毛一雷 孙 洁
（中国科学技术信息研究所，北京 100038）

摘要：当前科技查新工作的特点是高人力、低效率、难复制，查新结果的质量受查新人员业务水平和领域背景知识影响较大，纯粹依靠人工进行查新检索和对检索结果相关性判别无论是从效率还是准确率方面均无法适应科技创新对科技查新工作的新要求。在大数据时代，计算机技术和人工智能的介入可以在一定程度上提高查新的效率和质量。首先提出适用于科技查新业务的相关文档探测方法，将可用信息从文本相似度拓展到词汇、主题和语义维度，来捕捉查新点和科学技术要点与相关文档的关联关系，进而抽取相关特征并将其集成到条件随机场中进行相关文档探测。然后以全国科技查新事实型数据库为数据基础开展实验。实验表明，本文所提出的相关文档探测方法取得了较好的效果，有助于从数据科学和人工智能的角度来理解科技查新的业务和数据，为科技查新的自动化、智能化提供相应参考。

关键词：科技查新；相关文档探测；条件随机场；特征选取；文本相似度；共现词汇
中图分类号：G252.6 **文献标识码：**A **DOI：**10.3772/j.issn.1674−1544.2020.01.009

Research on the Application of Related Document Detection Method in Sci-tech Novelty Retrieval

CAO Yan, HE Xiaomin, CHEN Liang, MAO Yilei, SUN Jie
(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract: The current characteristics of sci-tech novelty retrieval are high manpower, low efficiency, and difficult to copy. The quality of the retrieval results is greatly influenced by the professional level and domain background of the search staff. Purely relying on manual search and correlation discrimination of search results cannot meet the new requirements of scientific and technological innovation in terms of efficiency and accuracy. In the era of big data, the involvement of computer technology and artificial intelligence can improve the efficiency and quality of new search to a certain extent. This paper proposes a method for detecting related documents that is suitable for new technology search business, extending the available information from text similarity to vocabulary, thematic and semantic dimensions to capture the relationship between the search points and scientific and technological points and related documents, and then extracts relevant features and integrates them into conditional random fields for related document detection. Experiments are carried out on the basis of the national science and technology novelty fact database. The experiments show that the document detection method proposed in this paper has achieved good results, which is helpful for understanding the new business and data of scientific and technological search from the perspective of data science and artificial

作者简介：曹燕（1980—），女，中国科学技术信息研究所副研究员，研究方向：技术经济与管理（通信作者）；**何晓敏**（1992—），女，中国科学技术信息研究所硕士研究生，研究方向：情报学；**陈亮**（1982—），男，中国科学技术信息研究所副研究员，研究方向：科学研究管理、图书情报与数字图书馆；**毛一雷**（1994—），女，中国科学技术信息研究所研究实习员，研究方向：情报学理论与方法；**孙洁**（1980—），女，中国科学技术信息研究所助理研究员，研究方向：科技查新。

基金项目：中信所重点工作项目“基于知识库的创新调查集成服务环境建设（一期）”（ZD2019−03）。
收稿时间：2019 年 11 月 27 日。

intelligence, and provides a corresponding reference for the new automation and intelligence of scientific and technological search.

Keywords: sci-tech novelty retrieval, related documents, conditional random field,feature selection, text similarity, co-occurrence vocabulary

作为科技情报工作的重要组成部分，科技查新是以反映查新项目主题内容的查新点为依据，以计算机检索为主要手段，以获取密切相关文献为检索目标。近年来，全国查新机构的查新项目整体数量保持上升趋势，部分机构的查新数量大幅度增加。在大数据时代，随着计算机和互联网的快速发展，中文数据库不断增多，不同领域技术间的关联度加强，复杂程度加深，多领域交叉地带已成为创新高发区。然而，科技查新工作主要是由人工来完成的，研究人员及查新人员很难在短时间内把握科技文献发展脉络及知识内容的关联[1]，且由于人员素质的参差不齐也可能会出现由于人的主观判断失误而对相关文献归类错误的情况。因此，单纯依靠人工进行查新检索和对检索结果相关性判别无论是从效率和准确率方面均无法适应科技创新对科技查新工作的要求，依靠计算机辅助手段帮助查新人员完成自动化文献检索和对比，可以提升查新工作本身的智能化水平，是未来科技查新数据挖掘研究中的重要问题。

1　国内外研究现状

笔者利用设定检索条件和人工去重筛选的方法，对《中国学术期刊（网络版）》的中国学术期刊全文数据库进行检索，获得 1989—2018 年有关科技查新的文献 3463 篇。对其进行高频词聚类分析，得到查新领域研究热点的变迁。根据分析结果，将科技查新领域的研究分为 3 个阶段（图 1）。第一阶段（1989—1999 年）研究以“查新工作实践”为主；第二阶段（2000—2009 年）注重对查新机构规范化和查新质量控制的研究；

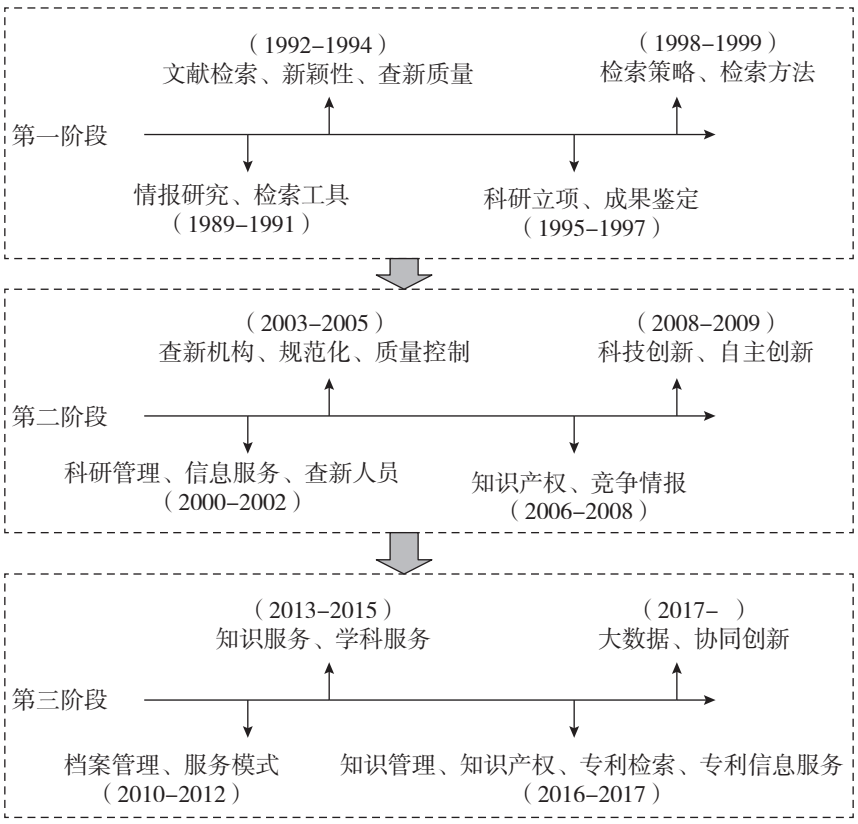

图 1　科技查新的 3 个研究阶段[2]

第三阶段（2010—至今）科技查新的基本要素如服务对象、查新流程、查新人员、检索方法与策略等研究已经发展到了一定规模，与此同时科技查新与大数据相结合，催生了科技查新信息化与智能化的研究与发展。

研究发现，以人工检索与判别为主的传统科技查新存在两个不足：一是严重依赖人力和专家资源，在科技发展速度不断加快，科技查新业务爆炸性增长的今天，传统方式难以应对海量业务；二是科技查新并非是简单的文献检索问题。文献检索应与情报分析相结合，将科技查新相关文献的判定过程转化为相关度的研究过程，即根据文献能够覆盖查新点的多少来确定文献与查新点的相关度[3]。随着人工智能技术的发展，利用机器学习技术更新和升级科技查新流程势在必行。

目前，国内外学者对科技查新工作已经开展了深入的研究，取得了一定的成果。武夷山[4]指出，科技查新似乎是中国独有的概念，在发达国家查新（Novelty Search）一般只与专利文献检索相联系，如瑞典专利注册局就有查新业务。另外有些国家还提供专利性检索的服务，由政府出面委托具有社会性质的咨询机构来完成，这些咨询机构相当于国内的查新机构，它们的服务模式与国内的科技查新服务相似[5]。在科技查新的业务流程中，检索环节对查新人员的专业性提出极高的要求，因此一些学者开始关注一些查新自动化、智能化的相关研究。王庆红等[6]设计了一种自助式查新方法，该方法可根据用户输入的查新内容进行语义分析，提取查新关键词，并通过关联检索生成相应的关联关键词，进而制定相关的检索式以完成查新点与数据库中相关文献的相关度比对。王培霞等[7]利用关键词抽取和领域特征扩展相结合的递进式迭代抽取方式实现了科技查新领域检索词的智能抽取。马林山等[8]设计了基于文档主题生成（LDA）模型的查新辅助分析系统设计功能框架，采用LDA模型通过相关文献关键词实现了潜在主题挖掘。这些研究都是从关键词入手检索相关文献，实现查新系统自动化的，而实际上用于判断相关文献和查新点是否相关的依据还有很多。相似度计算是信息处理领域中一项基础而又重要的工作，是实现文本数据挖掘的关键技术，许多重要的应用研究都与其有关。经过大量的文献调研，发现相似度的计算方法受到了研究人员的广泛关注。文本相似度的度量一直是文本挖掘领域的热点问题，优良的文档相似度算法可以实现对文档之间相似度的精准界定[9]。

在借鉴已有学者研究和算法的基础上，本文将提出一种基于图的语义模型，用来表示文档的内容，基于语料库和世界知识的方法，用一种应用语义相似度的探测方法来研究科技查新数据的深层含义；通过对科技查新报告中项目名称、技术要点、创新点、检索词进行主题概念抽取，构建领域知识库，探索适用于科技查新数据的算法。

2　数据源及特征选取

2.1　数据源

本文实验数据集来源于中国科学技术信息研究所构建的科技查新事实型数据库。该数据集共有2000条文本，其中包含的字段有：查新报告id、查新报告编号、查新点、检索式、项目名称、委托人、委托日期、查新目的、所属技术领域、学科领域、查新结果、查新结论、查新项目的科学技术要点、查新范围、检索范围、文献检索范围及检索策略。涉及的学科领域有地球科学、测绘科学技术、生物学、动力与电气工程、化学工程、土木建筑工程、电子通信与自动控制技术、计算机科学技术、材料科学、机械工程、临床医学等。

2.2　特征选取

本文在充分考虑查新点和项目科学技术要点本身的状态特征，以及查新点与项目的科学技术要点与相关文献之间的相互关系即转移特征的基础上，选取了文本相似度、共词主题聚类、共现词汇特征、共现词汇数目等特征。

2.2.1　文本相似度

本文采用的是语义相似度计算方法，在语义相似度计算方法的研究方面已经有了很多成果，

包括基于字面匹配的语义相似度计算方法、基于潜在语义分析的概率主题语义相似度计算方法和基于深度学习的语义相似度计算方法。本文的相似度计算方法是在TF-IDF的基础上采用潜在语义索引（LSI）模型[10]。

文本相似度的计算流程如图2所示。

2.2.2　共词主题聚类

共词主题聚类属于定性分析法，只是对共词数目进行频次统计，还不能充分揭示科技查新数据的内在关联，难以反映科技查新数据的语义信息状况[12]。主题提取是一项基础性的信息提取工作，主题聚类则是以主题提取为前提的信息聚类过程。共词主题聚类就是对查新点和项目科学技术要点与相关文献的两两共现的词语进行聚类分析。

对查新点和项目科学技术要点与相关文献进行聚类主要包括三个步骤：首先是将共现词汇进行多文本的预处理，构建特征值的向量空间；然后通过上一步构建的向量空间继续构建相似性矩阵；最后在聚类算法中加入样本密度排序对词汇进行排序，根据相似性矩阵生成的簇的数目来确定聚类的主题数。

本实验使用的是LDA聚类算法，当输入一个文档的集合$D=\{d_1,d_2,d_3,\ldots,d_n\}$，同时需要输入聚类的类别数量m，然后得出每一篇文档的d_i在所有Topic上的一个概率值P，这样每一篇文档都可以得到一个概率的集合$d_i=\left(d_{p_1},d_{p_2},\ldots,d_{p_m}\right)$，并且同样的文档中的所有词也会得到其对应的每个Topic的概率，可以记为$w_i=\left(w_{p_1},w_{p_2},\ldots,w_{p_m}\right)$。这样就得到了两个矩阵：一个是文档的Topic矩阵，一个是词语的Topic矩阵。通过上面所描述的算法，将文档和文档中的词语投射到一组Topic上，这样就可以通过这组Topic找到文档与文档之间、词语与词语之间、文档与词语之间的潜在语义关系。由于LDA属于无监督的算法，每个Topic并不会要求设定条件，但是通过聚类后，可以统计出各个Topic上词语的概率分布，那些在Topic上并且概率较高的词语就可以较好地描述此Topic的含义。

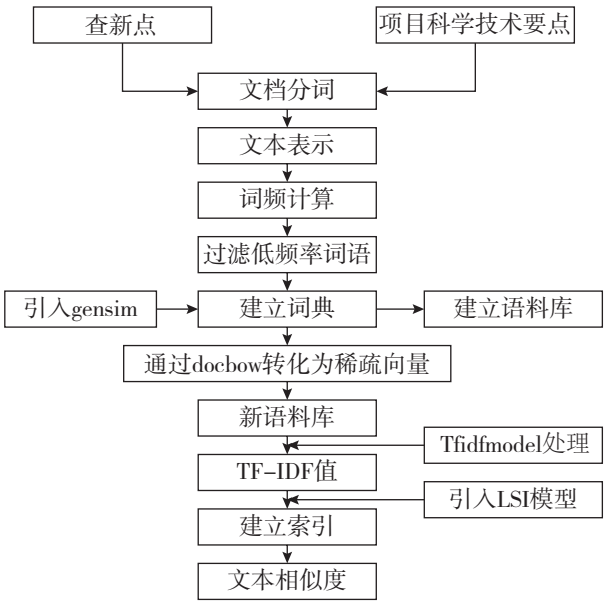

图2　文本相似度计算流程

2.2.3　共现词汇特征

本实验使用信息增益方法来度量查新点和项目科学技术要点与相关文献共现词汇的重要程度，将去重后的词汇按照信息增益的大小进行排序。

信息增益（information gain）表示得知特征X的信息而使得类Y的信息的不确定性减少的程度[11]。特征B对训练数据集C的信息增益为$g(C,B)$，是集合C的经验熵$H(C)$与特征B在给定条件下D的经验条件熵$H(C\mid B)$之差，即：

$$g(C,B)=H(C)-H(C\mid B)$$

其中，熵是表示随机变量不确定性的度量，熵越大，随机变量的不确定性就越大。条件熵$H(C\mid B)$表示在已知随机变量B的条件下随机变量C的不确定性。当熵和条件熵中的概率由数据估计得到的时候，此时的熵称为经验熵，而条件熵则称为经验条件熵。此时，如果有0概率，令$0\log 0=0$。另外，一般的熵与条件熵之间的差值称为互信息。在决策树学习中，信息增益等价于训练数据集中类与特征的互信息。

2.2.4　共现词汇数目

共现词汇数目的原理比较简单，是最简单的评估函数。它取的是查新点和项目技术要点与所有相关文献的交集。由于共现词汇数目的取值范

围为[0，+∞)，共词数目的特征的方差过大，使参数估计器无法正确地学习其他的特征。所以需要对共现词汇数目进行标准化，将共现词汇数目缩小到一个指定的范围。

本文采用的是sklearn.preprocessing中的scale方法，使特征可以分布在一个最大值和最小值之间，通过MinMaxScaler或者MaxAbsScaler方法可以使此维标准分布在[−1，+1]之间。

线性归一化（MinMaxScaler）为min−max标准化，也称为离差标准化，是对原始数据的线性变换，min−max标准化方法的缺陷是当有新的数据加入时，可能会导致Xmax和Xmin的值发生变化，是需要重新计算的。

$$X_{scaled}=\frac{X-X_{\min(axis=0)}}{X_{\max(axis=0)}-X_{\min(axis=0)}}\times(\max-\min)+\min \quad (1)$$

其中，max和min是给定放缩范围的最大值和最小值。

3　基于CRF的相关文档探测方法的应用

根据科技查新数据的特征和条件随机场（Conditional Random Field, CRF）模型的使用范围，来构建适用于科技查新数据的条件随机场模型。图3为相关文档探测方法的总体流程设计。首先对初始数据集进行语料标注，对语料进行分词、停用词与文本表示等预处理，得到实验要用的数据集。然后对数据集分别提取文本相似度、共词主题聚类、共现词汇特征、共现词汇数目等特征。最后将数据集按7:3的比例分为训练集与测试集，并将每条数据集的特征集成到基于CRF的科技查新相关文档探测模型中。

3.1　适用于科技查新数据的图结构

CRF是在给定一组输入随机变量条件下另外一组输出随机变量的条件概率分布模型，是一种判别式的概率无向图模型。既然是判别式，那就要对条件概率分布建模，其特点是假设输出随机变量构成马尔科夫随机场[12]。根据科技查新的工作流程与具体的数据特征形式，查新人员在进行相关文献的探索时可将过程用图4表示，得出查新点和项目科学技术要点与相关文献之间的关系形式。

查新点或项目科学技术要点与相关文献之间的关系可以用条件随机场的形式表示，如式（2）所示，其中X代表观测出来的所有特征，包括单

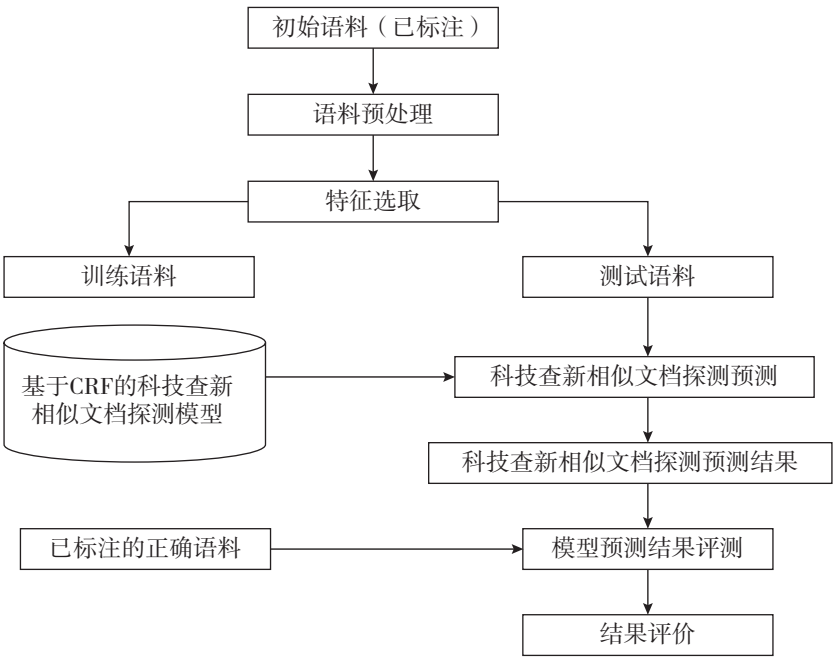

图3　基于CRF的相关文档探测方法

独结点的特征和表示两个结点之间关系的特征，Y代表两个节点之间是否具有相关关系。

$$P(Y|X)=\frac{\mathrm{P}(Y,X)}{P(X)}=\frac{\mathrm{P}(Y,X)}{\sum_Y \mathrm{P}(Y,X)} \quad (2)$$

图 4 可以简化为以下 3 种特征关系。

（1）查新点和科学技术要点本身的状态特征，如查新点所在的技术领域标签，即为$f(x_i)$，i=1, 2, …, n。

（2）查新点和科学技术要点与相关文献之间的关联特征，称为状态特征，如文本相似度、共现词，记为$f(x_i, x_j)$，i=1, 2, …, n，j=1, 2, …, m。

（3）$b_1, b_2, \ldots, b_n$由于连接同一查新点而产生的共现特征，称为转移特征记为$f(y_k, y_s)$，k =1, 2, …, m，s =1, 2, …, m。

所以：

$$P(Y,X)=e^{f(x_i)+f(x_i,\ y_j)+f(y_k,\ y_s)} \quad (3)$$

$$\sum_Y P(Y,X)=\left(\sum_Y e^{f(y_j)+f(y_k,\ y_s)}\right)e^{f(x_i)} \quad (4)$$

得到：

$$P(Y\mid X)=\frac{e^{f(x_i)+f(x_i,y_j)+f(y_k,y_s)}}{\left(\sum_Y e^{f(y_j)+f(y_k,y_s)}\right)e^{f(x_i)}}=\frac{e^{f(x_i,y_j)+f(y_k,y_s)}}{\left(\sum_Y e^{f(y_j)+f(y_k,y_s)}\right)} \quad (5)$$

3.2　条件随机场模型的简化

当引入相关关系之间的转移特征时，如图 5，仅 11 个相关关系节点就使得团结构形式非常复杂。难以精确求参，只能采用近似求参的方法。

经过实验，发现采用近似求参的CRF比精确求参的CRF低约 13 个百分点，比只引入状态特征的CRF低约 8 个百分点。所以本实验将相关关系之间的转移特征舍去。

最终化简后的模型为：

$$P(Y\mid X)=\frac{e^{f(x_i)+f(x_i,y_j)+f(y_k,y_s)}}{\left(\sum_Y e^{f(y_j)+f(y_k,y_s)}\right)e^{f(x_i)}}=e^{f(x_i,y_j)} \quad (6)$$

代入各个特征函数可以得到：

$$P(Y\mid X)=\frac{e^{w_1 f_1(x_i,y_j)+w_2 f_2(x_i,y_j)+w_3 f_3(x_i,y_j)+w_4 f_4(x_i,y_j)}}{\sum_Y e^{w_1 f_1(x_i,y_j)+w_2 f_2(x_i,y_j)+w_3 f_3(x_i,y_j)+w_4 f_4(x_i,y_j)}} \quad (7)$$

其中，特征函数$f_1(x_i,y_j)$代表的特征是文本相似度，$f_2(x_i,y_j)$代表的特征是共词主题聚类，$f_3(x_i,y_j)$代表的特征是共现词汇特征，$f_4(x_i,y_j)$代表的是共现词汇数目。w_1, w_2, w_3, w_4表示各个特征在分类中的权重大小。

4　结果分析

本文从有效性和复杂性两个方面对模型进行评价[14]。有效性主要是指在现有的模型算法下，文本分类结果的准确性；复杂性则是从时间和空间两个角度出发，涉及文本分类过程中存储空间大小、模型运行过程中所消耗的时间等因素。本文有效性的评价指标有准确率（Accuracy）、精确

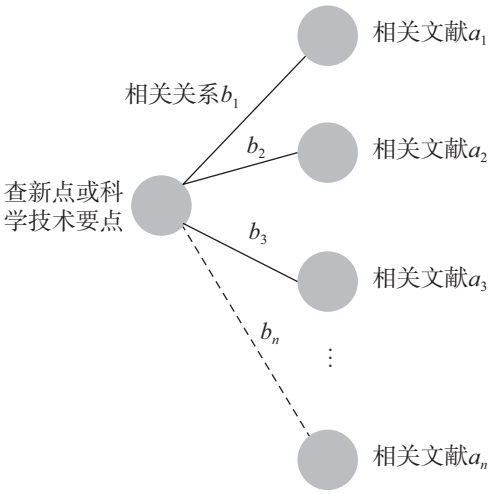

图 4　科技查新数据图结构

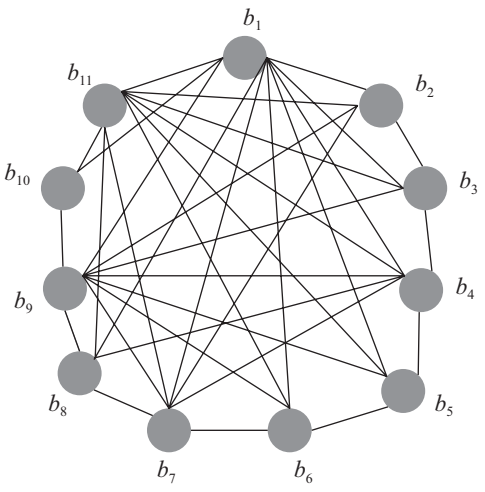

图 5　相关关系团结构

率（Precision）、召回率（Recall）和综合评价指标F_1值（对精确率和召回率的综合评价）（表1）。

实验经过多次调整参数，选择参数C=20为模型性能最好的情况，C为超参数，指的是正则系数，一般在log域（取log后的值）。在不同特征数目的条件下，准确率、精确率、召回率和综合评价指标F_1值如图6所示。

从表1和图6可以看出：加入共现词汇特征后的条件随机场模型在精确率、召回率、F_1值上都比没有加入词汇特征时表现得更有优势，效果更好。

当只用文本相似度、共现词汇数目、共现词汇主题分类以及C=10时进行训练模型时，精确率、召回率、F_1值都比较低。若把参数C调整到20时，模型的指标提升的幅度非常小。当在模型中加入50个特征词汇并且将参数C设为20时，精确率、召回率、F_1值都有明显上升。随着特征数由50到450，除了精确率的变化幅度较小（基本上稳定在87%），召回率和F_1值随着特征数目的增多而有了显著提高。当特征数到达500并且高于500时，精确率、召回率、F_1值都开始下降。

由此可见，当用查新点与相关文献进行相关性探测时，准确率几乎没有什么变化，基本稳定在99%以上；精确度并不随着特征数目的增加而提升，其最好的状态是在450特征时，达到了88.52%；召回率是随着特征数目的增加而逐步增加，在450特征时，达到了70.39%；F_1值也是随着特征数目的增加而逐步增加，在450特征时，达到了77.62%。

5　结论

在科技查新业务工作中，文献检索和对检索文献分析对比是科技查新进行新颖性判断的重中之重，但是目前的科技查新业务主要是依靠人工进行，处于一种高人力、低效率、难复制的困境。在大数据时代，计算机和人工智能介入科技查新是大势所趋，用自动化手段帮助查新人员开展查新工作在一定程度上可以提高查新的效率和质量。本文构建了一种基于多维度特征的相关文档探测方法，来捕捉查新点相关文档的关联关系，并将相关特征集成到条件随机场中进行相关文档探测。实验发现，在科技查新领域中，文本

表1　评价指标

指标	说明
准确率	指所有的预测正确的占总的比重
精确率	指正确预测为正的占全部预测为正的比例
召回率	指正确预测为正的占全部实际为正的比例
F_1	指精确率和召回率的综合评价

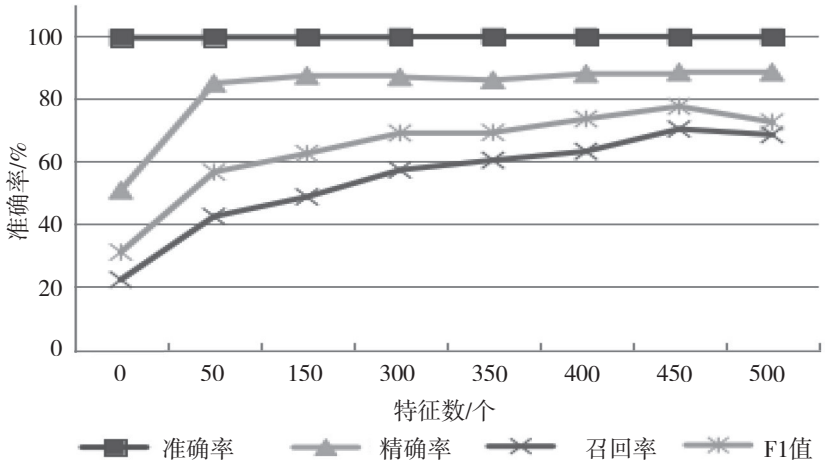

图6　查新点性能评价指标折线图

相似度特征在模型中的贡献较弱，贡献度较大的是适当数目的共现词汇，间接表明了文本相似度这一指标只是科技查新要考虑的诸多因素的一种，除此之外还有很多因素，可以对科技查新效果提供直接贡献，如共现词汇及其数量、专业词汇、同义词表等，这说明要提升科技查新效果，领域词表的构建至关重要。此外，可以引入其他信息类型，如文献类别、引文信息、作者信息等，这些都能够提升科技查新效果。

参考文献

[1] MEZA B A. Searching and ranking documents based on semantic relationships[C]// Proceedings of the 22nd international conference on data engineering workshops. Washington: IEEE Computer Society, 2006: 1060−1064.

[2] 何晓敏, 曹燕, 陈亮, 等.1989—2018年科技查新内涵的文献计量分析[J].中国科技资源导刊, 2018, 50(6): 55−62.DOI: 10.3772/j.issn.1674−1544.2018.06, 009.

[3] 罗靖, 王立新, 杜正飞.论云南省科技查新事实型数据库建设[J]. 图书情报工作, 2014(S2): 63−65.

[4] 武夷山.从科技查新向广义的专业化信息检索服务转型[J].情报学报, 2014, 33(7): 前插1.

[5] 陈守鹏. 科技查新工作中相关文献的认定与筛选[J]. 现代情报, 2009, 29(12): 145−147.

[6] 王庆红, 韦嵘晖, 等.一种自助式查新系统与方法: 中国, 201510696381.6[P].2016−03−09.

[7] 王培霞, 余海, 陈力. 科技查新中检索词智能抽取系统的设计与实现[J]. 现代图书情报技术, 2016(11): 85−96.

[8] 马林山, 郭磊. 基于主题模型(LDA)的查新辅助分析系统设计研究[J]. 现代情报, 2018(2): 111−115.

[9] 邵媛媛. 基于语义的文本相似度算法研究[J]. 计算机光盘软件与应用, 2014, 17(9): 302−303.

[10] 章成志.主题聚类及其应用研究[M]. 北京: 国家图书馆出版社, 2013.

[11] 王刘阳.文本分类中特征选择与加权算法的研究[D]. 杭州: 杭州电子科技大学, 2016.

[12] 李航. 统计学习方法[M]. 北京: 清华大学出版社, 2012: 191−205.

（上接第53页）

[8] 杨爽.企业技术创新的信息保障体系构建研究[J].情报科学, 2012, 30(3): 460−463.

[9] 房惠玲.甘肃为战略性新兴产业建免费开放专利数据库[N].甘肃经济日报, 2014−11−14(1).

[10] 郑祖婷, 史宝娟.唐山市构建战略性新兴产业信息服务平台研究[J].科技和产业, 2012, 12(6): 14−16.

[11] 霍国庆, 李天琪, 张晓东.战略性新兴产业信息资源保障体系建设[J].重庆社会科学, 2012(6): 79−85.

[12] 贺正楚, 吴艳.战略性新兴产业信息资源服务体系与网络服务平台研究[J].中国科技论坛, 2013(4): 21−27.

[13] 黄威. 美国下一代广播电视网发展的基础: 国家宽带计划的政策内涵及措施[J].有线电视技术, 2013, 20(7): 6−10.

[14] GALE A Boyd. A Method for measuring the efficiency gap between average and best practice energy use: The ENERGY STAR industrial energy performance indicator [J]. Journal of Industrial Ecology, 2005(3): 15.

[15] 邓胜利, 周婷.面向战略性新兴产业的信息服务与保障研究[J].情报理论与实践, 2012, 35(7): 6−8, 13.

[16] 霍国庆, 王少永, 孙皓.我国战略性新兴产业共性信息资源需求的实证研究[J].信息资源管理学报, 2015, 5(1): 4−11.

[17] 叶继元, 谢欢.存量与增量: 中国战略性新兴产业信息资源保障体系的宏观思考[J].图书馆论坛, 2015(1): 1−7.

[18] 潘玉辰.基于大数据下战略性新兴产业个性化信息资源服务模式研究[J].开发研究, 2016(3): 20−25.

[19] 孙振, 郑德俊.面向战略性新兴产业的信息服务模式选择[J].情报科学, 2014, 32(4): 68−71, 100.

[20] MENG X. Research on coordination development between electronic commerce and modern service industry[J].Journal of Computers, 2011, 6(7): 1461−1468.

[21] 张晓东, 霍国庆.战略性新兴产业信息资源服务模式与竞争力分析[J].科技进步与对策, 2013, 30(2): 74−78.

[22] 孙振.战略性新兴产业信息服务模式研究[D].南京: 南京农业大学, 2014.

[23] 刘珺.战略性新兴产业的信息资源服务创新研究[J].科学管理研究, 2015, 33(6): 52−55.