

一种参考文献自动比对方案

包震宇

上海师范大学期刊社, 200234, 上海

摘要 针对工科类论文参考文献著录情况较为混乱的现状, 本文提出一套参考文献自动比对方案, 简化了参考文献的检查工作。利用 C# 编程语言, 设计了文献信息系统和文献比对系统, 将作者原稿录入文献信息系统, 并与文献比对系统中的规范文献库进行比较, 最终的参考文献规范版本将显示在原稿中。通过设计文献信息系统, 降低了参考文献比对的人力、时间及经济成本, 有效提高了参考文献著录的准确性及效率。

关键词 文献信息系统; 文献比对系统; 参考文献; 规范化

An automatic comparison scheme of references//BAO Zhenyu

Abstract In view of the chaotic situation of the bibliographic description of engineering papers, a set of reference standardization scheme is proposed to simplify the process of editing, proofreading and the inspection of references. Using C# language programming, we design the reference information system and the reference comparison system. We input the references into the reference information system, and compare the references in the author's original with the standard literature database in the system to reduce the errors in the description of references in the original. Then we can display the standard bibliographic description in the original. Through the design of reference information system, the manpower, time, and economic cost of reference comparison are reduced, and the accuracy and efficiency of reference description are effectively improved.

Keywords reference information system; reference comparison system; references description; standardization

Author's address Academic Journals Center of Shanghai Normal University, No. 100 Guilin Road, Xuhui District, 200234, Shanghai, China

DOI: 10.16811/j.cnki.1001-4314.2020.05.013

工科类论文的参考文献大多来源于国外学术刊物或学术会议, 没有统一的格式, 以致国内作者在引用文献的过程中, 参考文献的著录方式与我国指定的标准相去甚远^[1-6]。然而将作者原稿中的参考文献格式修改到符合国家标准程度, 工作量大, 难度高, 并且容易出错^[7-12]。如何有效提高工科类论文的参考文献编校工作, 成为一个亟待解决的问题。

针对上述问题, 研究人员提出了一些有效的对策。侯修洲等^[13]提出基于 CrossRef 数据库的参考文献自动加工, 提升了参考文献编校工作的效率和准确率; 赵茜^[14]提出规则法、对照法、常识法、质疑法等 4 种行之有效的校对方法, 提高编辑工作效率, 减少参考文献著

录差错; 于学玲等^[15]提出科技期刊编辑应利用参考文献检索技能、专业知识挖掘参考文献的隐蔽性差错; 郑晓梅^[16]将科技期刊参考文献著录不规范的情况进行分类, 找出导致参考文献著录不规范或差错的主要原因, 制定确保参考文献著录规范化、准确化的应对措施。以上方法在实施过程中, 严重依赖作者对参考文献著录标准的理解程度, 对减轻参考文献的编校工作量作用有限。

为此, 笔者利用 C# 编程语言, 设计了文献信息系统和文献比对系统, 将原稿中的参考文献输入信息系统, 在规范文献库中检索相应的条目, 并替代原稿中的参考文献。

1 规范文献库的设计与实现

规范文献库中的文献条目来源于 IEEE, 由于国外数据库中的文献条目没有固定格式, 也不符合 GB/T 7714—2015《信息与文献 参考文献著录规则》的规定, 在录入文献库之前, 需要对原始文献条目进行预处理。

1.1 文献条目预处理

以 IEEE 平台下的《IEEE Transaction on Communication》杂志为例, 文献责任者在其官网的显示格式如图 1 所示。

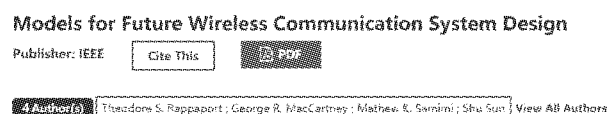


图1 IEEE 原始文献责任者信息

对于主要责任者, GB/T 7714—2015 规定其著录格式必须符合以下规则:

- 1) 个人著录采用姓在前名在后的著录形式;
- 2) 欧美著者的名可用缩写字母, 缩写名后省略缩写字点;
- 3) 用汉语拼音书写的人名, 姓全大写, 其名可缩写, 取每个汉字拼音的首字母;
- 4) 主要责任者不超过 3 个时, 全部照录; 超过 3 个时, 著录前 3 个责任者, 其后加“等”或与之相应的词(英文文献用“et al”)。

因此,对于图1中的主要责任者信息,应在录入规范文献库之前,将其转换为“RAPPAPORT T S, MACCARTNEY G R, SAMIMI M K, et al”。

另外,为实现全刊统一,还须遵循 GB/T 7714—2015 中未强制规定的一些规则,如:英文文献题名中除句首字母外,其余字母小写;出版者除介词外,所有单词首字母大写,卷号与期号不可同时缺失等。按照上述规则,设计文献信息系统的数据库预处理模块,使其能够自动识别并转换原始数据。

1.2 数据校验

除文献来源平台的原始数据不规范之外,作者在撰写参考文献的过程中,可能也会因某些疏忽,录入了错误的内容,如:文献条目的页码为“123-144”,作者在复制页码信息时少了一位页码数,变为“123-14”,导致该文献的起始页码大于终止页码。针对如此明显的差错,笔者设计了文献信息系统的数据库校验模块,使文献信息系统能够识别一些显而易见的差错,并进行最大程度的修正。设计的纠错规则如下:

- 1) 主要责任者姓名中不出现数字;
- 2) 起始页码不可大于终止页码;
- 3) 转页码不可小于终止页码;
- 4) 出版时间不可大于文献录入时间。

1.3 流程设计与实现

规范文献库的生成流程如图2所示。

首先,判断参考文献所有必需的信息是否有缺失,并判断参考文献责任人姓氏与名字是否成对出现,满足上述条件后,依照“姓前名后”的原则,将英文姓名全改为大写,名字只取每个单词的首字母,以空格分隔;然后,将题名信息首字母改为大写,并将出版者信息每个单词首字母改为大写;接着,判断录入信息是否满足期号、卷号、起始页码及终止页码的约束条件;最后生成最终的参考文献条目,并导入规范文献库,通过 Excel 文件查看。

采用 C# 编程语言实现文献信息系统的的功能,文献录入界面如图3所示。

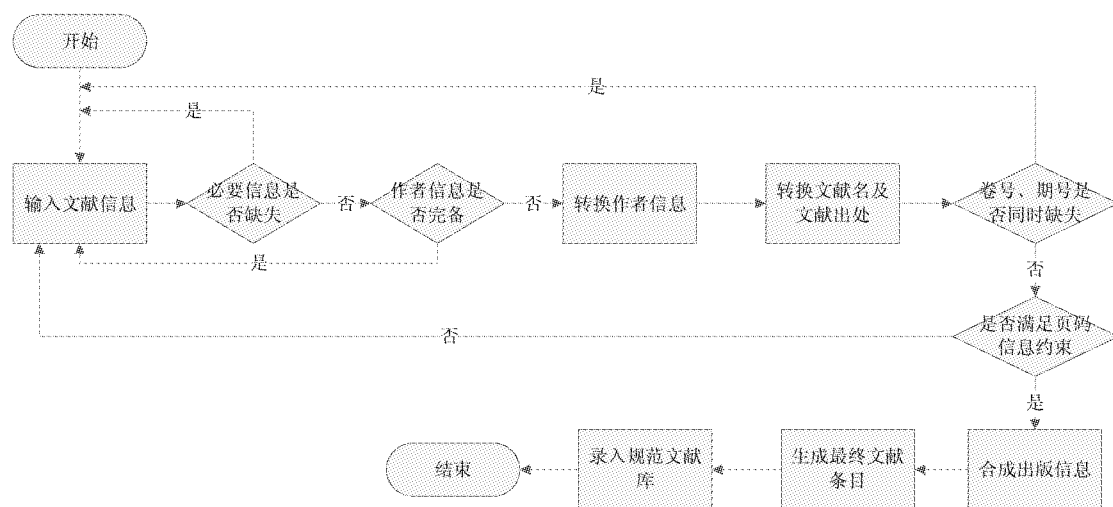


图2 生成规范文献库



图3 规范文献库录入界面

2 参考文献比对

为了使用不同的文献题名来识别原稿中的参考文献信息,笔者设计了一套文献比对系统,并替换原稿中的参考文献,流程如图4所示。

首先手动输入单条原稿中的参考文献题名,由系统进行比对,并在数据窗口中显示规范文献库中的文献条目,再手动输入下一条参考文献进行比对,生成规范文献条目,以此循环执行,直到所有原稿中的参考文献条目被输入一遍,得到与之对应的文献格式的规范表述,最后导出至 Excel 文件中,以便将最终的参考文献内容复制到原稿中。

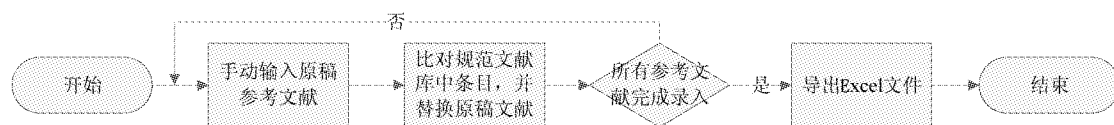


图4 文献比对系统流程

在文献比对系统的执行过程中,比对程序的设计是核心部分。为防止因英文文献大小写差异、题名字母缺失,或者题名单词之间空格缺失而造成无法查询规范文献库中相应条目的问题,系统执行的是模糊查询操作,遵循“多查不漏”的原则,在确保查询到规范文献条目的前提下,检索最少量的文献条目,最终由编辑选择正确的文献条目,删除其余文献条目。

在整个文献比对的过程中,编辑参与的工作只有手工输入原稿的参考文献以及最后将 Excel 文件中的参考文献复制到原稿中,省去了对参考文献格式及内容检查这项繁杂的工作,系统可确保文献的最终著录格式及内容的准确性,减少了参考文献编校的工作量。

综上所述,笔者设计的这套参考文献处理方案具备以下优势:

1) 2个文献处理系统实施参考文献的编校工作,减少了由人为因素引起的差错,提高了参考文献的编校效率,降低了人力成本;

2) 2个文献处理系统复用性较高,且便于维护,相对于人工编校流程,更为简便、快捷,降低了时间成本;

3) 市场上的文献检测服务一般在2元/条左右,笔者设计的方案兼容性较好,可在任何 Windows 操作系统下免费运行,降低了经济成本。

3 结束语

提出了一套参考文献处理方案,利用 C#编程语言,设计了文献信息系统及文献比对系统,对原稿中的参考文献逐一进行比对、替换。文献信息系统用以生成规范文献库,为后续规范原稿的参考文献提供依据;文献比对系统在规范文献库中逐一检索与原稿的参考文献相对应的条目,并实施选择性替换工作。结果表明:该套方案在一定程度上能够简化参考文献的编校工作,减少作者原稿中出现的差错数量。但是方案仍存在问题有待优化,如:对参考文献的比对严重依赖于规范文献库,一旦规范文献库的准确性得不到保

证,方案的实施效果将大打折扣;方案还无法对析出参考文献的表述进行规范处理,这些都将在后续的研究工作中加以改进。

4 参考文献

- [1] 祝青松. 基于 GB/T 7714 研究工作中的参考文献著录趋势分析[J]. 编辑学报, 2016, 28(4): 352
- [2] 陶范. 参考文献引用原则辨析[J]. 编辑学报, 2006, 18(4): 254
- [3] 曹敏. GB/T 7714—2015《信息与文献 参考文献著录规则》标准解析[J]. 科技与出版, 2015(9): 41
- [4] 程晓英. 科技论文参考文献的核实、规范与优选[J]. 科技与出版, 2015(8): 51
- [5] 马建华. 学术论文参考文献的隐形错误例解[J]. 编辑学报, 2019, 31(2): 169
- [6] 丁春. 论参考文献的引用原则、著录要求及编辑审稿要点[J]. 中国科技期刊研究, 2008, 19(4): 534
- [7] 陈浩元. GB/T 7714 新标准对旧标准的主要修改及实施要点提示[J]. 编辑学报, 2015, 27(4): 343
- [8] 周园. 著录外文参考文献值得注意的几个问题[J]. 中国科技期刊研究, 2008, 19(3): 488
- [9] 王祥国. 参考文献著录需注意的5类特例[J]. 编辑学报, 2016, 28(1): 24
- [10] 刘永强, 徐敏, 李园, 等. 一种参考文献辅助比对方法[J]. 编辑学报, 2019, 31(4): 431
- [11] 谢金海. 编辑应引导作者重视参考文献著录[J]. 编辑学报, 2012, 24(1): 47
- [12] 张宏, 李航, 程立冬, 等. 运用 EndNote 批量编辑加工英文参考文献[J]. 编辑学报, 2018, 30(4): 369
- [13] 侯修洲, 黄延红. 基于 CrossRef 数据库的参考文献自动加工及 XML 标引方法[J]. 编辑学报, 2017, 29(1): 70
- [14] 赵茜. 科技论文参考文献的校对方法[J]. 编辑学报, 2009, 21(5): 416
- [15] 于学玲, 阎明凡. 科技期刊编辑要善于发现参考文献的隐蔽性差错[J]. 编辑学报, 2014, 26(2): 132
- [16] 郑晓梅. 正确著录参考文献:编辑的事前作者工作[J]. 编辑学报, 2014, 26(5): 441

(2020-05-22收稿;2020-07-17修回)