

学术期刊虚假网站特征与投稿官网查询的软件实现*

晁晓筠 李德根[†]

黑龙江科技大学学术理论研究部,150022,哈尔滨

摘要 学术期刊虚假网站侵害了作者和出版者的合法权益,损害了学术期刊的声誉。为应对学术期刊虚假网站的侵权行为,我们在分析虚假网站特征的基础上,提出防范假冒网站侵害的措施。所研制的学术期刊网址查询系统,通过定期维护更新,增强识别虚假网站的便捷性,使作者可以实时查找到学术期刊的真实网站,期望有助于减少因虚假学术期刊网站泛滥给作者、出版者造成的困扰,为维护学术期刊网站的网络运行环境提供一种技术防范措施。

关键词 学术期刊;查询系统;虚假网站;防范措施

Features of fake websites of academic journals and the query software for official website's query// CHAO Xiaoyun, LI Degen

Abstract The fake websites of academic journals have infringed the legitimate rights and interests of authors and publishers and tarnished the reputation of academic journals. Based on analyzing the characteristics of false websites, we put forward measures to prevent the infringement of false websites for the purpose to cope with the infringement of false websites of academic journals. The website inquiry system is specially developed for academic journals, which is subjected to regular maintenance and updating, and allows potential authors to find websites of academic journals in real time. The inquiry software system is specially designed for web site finding, which serves the function of enhancing the convenience of identifying false websites and reducing the impact of false websites. These preventive measures and software technology hold the promise of reducing the overflow of false academic journals websites and thereby maintaining a better network environment for academic journal websites

Keywords academic journals; inquiry system; false websites; preventive measures

Authors' address Department of Academic Research of Heilongjiang University of Science and Technology, 150022, Harbin, China

DOI:10.16811/j.cnki.1001-4314.2019.04.024

目前,绝大部分学术期刊在征稿过程中,采用的都是在线投稿。据相关统计,到2015年底为止,被中国科学引文数据库(CSCD)收录的学术期刊中,采用网络在线投稿系统的期刊大约有73.3%,且绝大多数学术期刊将稿件的处理系统放置在其官方网站上。网站

上有明确的投稿模板且适时更新^[1]。然而,作者在搜索页面上输入相应期刊名称后发现,排列在前面的往往不是期刊的真实网站^[2]。

由于网络的监管力度不够、搜索引擎企业社会责任的缺失等诸多原因,在利益的驱动下,学术期刊的官方网站被“山寨”现象日渐严重。有些虚假网站利用技术的优势,竟比真正的网站更加符合学者投稿时的心理,致使很多作者上当受骗,损害了作者的利益。虽然部分作者进行了维权,但是最后的结果大多数因为投诉成本或者担心影响自己的学术声誉而放弃。由此,学者们对假冒网站侵权所造成的不良影响进行了研究。孙昌明等^[3]着眼于医学核心期刊,分析百度搜索引擎中的竞价业务对医学核心期刊的影响,发现我国中文核心医学期刊在百度竞价排名中深受其害,需要政府、期刊、作者共同努力维护自身权益和社会和谐。黄锋等^[4]在分析科技期刊假冒网站特征的前提下,提出识别假冒网站的方法和技巧,以维护作者的合法权益。黄仲一等^[5]分析了科技期刊假冒网站的特征,并阐述了假冒网站存在的原因,认为科研、评价体制不合理导致论文发表供需不平衡是假冒网站存在的主要原因,并就此提出了应对策略。杨继涛等^[6]采用了调查统计法,归纳总结了科技期刊遭遇网络侵权的类别,提出科技期刊应对网络侵权的对策。张颖^[7]介绍了非法期刊的发行渠道、特点以及辨别非法期刊的方法,帮助高校教师、科研人员提高防范意识。贺嫁姿等^[8]运用调查研究、网络搜索、案例分析等方法,分析论文中介网站冒名欺骗期刊作者的危害,提出建立期刊网站,使之容易被搜索引擎搜索。学者们的研究为遏制假冒网站扩张,维护学术出版秩序做出了重要的贡献。笔者在分析虚假网站特征基础上,开发投稿系统网址查询软件,为作者提供一种查询工具,运用技术措施减少虚假网站的侵害。

1 虚假网站的特征与辨识

学术期刊虚假网站运营的目的不是传播,而是敛财。为了更好地维护学者学术成果的发表权,识别学术期刊虚假网站显得尤为重要。笔者通过辨识学术期刊虚假网站的特点,帮助作者在投稿时提高警惕性。

1.1 虚假网站的特征 学术期刊虚假网站主要是模

* 黑龙江省高等教育教学改革项目(SJGY20180459)

[†] 通信作者

仿期刊真实官方网站的界面进行操作,以达到其大量收敛钱财的最终目的。其主要特征有以下3个。

1) 学术期刊虚假网站借助搜索引擎(如百度、搜狐、好搜等等)的相关、相近词语搜索频繁出现以引人关注。由于某种竞拍机制,在搜索引擎界面上,虚假网站可较多地出现在真实网站之前。当读者点击虚假网站之后,出现的界面里面通常有以下几种文字信息:一是突出文章处理时间快,一般都有诸如“3天之内完成审稿”“一个星期之内出结果”“优先安排出版”“版面抢排”“周期短”“见刊快”等提示;二是强调服务的完成率,通常有诸如“双方签合同”“一对一专家审稿服务”“资深专家进行文章修改”“杂志社在线征稿”等等;三是带有在线咨询的滚动条,当读者点击之后,自动弹出对话框,客服会主动询问您的需求,然后因势利导,一步一步诱骗作者,最后根据不同的学术期刊服务费用不同明码标价。笔者对相关虚假网站进行调查获得其版面费收取信息,如4000字左右的文章,发表在省级学术期刊,为800~1500元,北大中文核心期刊为8000元左右,南大核心期刊(CSSCI)为1万~2万元。其中,统计源期刊保证在付款后3个月内见刊。

2) 学术期刊虚假网站通过技术操作盗用正规学术期刊的相关信息,尤其是一些知名的刊物,虚假网站通过截取正规期刊刊名的首字母作为域名,其他内容与正规期刊的官方网站一样,也包含往期目录、期刊主办单位等信息。

3) 学术期刊虚假网站在要求作者汇款时的特点。合法的正规期刊在经过审核决定稿件录用后,通知作者缴纳版面费时,一般都是要求汇款至期刊的对公账户,如某某某期刊编辑部(收),而不会出现要求作者将版面费汇至私人账户的情况。学术期刊虚假网站则恰恰相反,其收款方一般为某个学术期刊+个人名字,而且个人账户还不止一个,通常为3个。

1.2 虚假网站的甄别 目前,学术期刊虚假网站技术的专业性和存在的隐蔽性,使作者难以直观对其进行甄别。掌握一些方法,可防范虚假网站的影响。一是观察域名。观察域名时要仔细,学术期刊的网站多数是建立在教育网或者政府网上,国家对这2个域名管理较为严格,冒充的可能性极低。二是浏览网页内容。可以通过网站网页所显示的内容去进行甄别。虚假网站的信息更新不及时,网页会有广告链接和提供帮助的对话框,联系方式大部分是QQ号或者微信号,很少是固定电话,即使是固定电话,也是打不通的空号。三是核实备案信息。真正的学术期刊网站会在工信部网站上进行备案,如果能查询到该网站的备案信息并且这些信息与期刊信息一致,则该网站为真实网站。反

之,则为虚假网站。四是辨别ISSN号。通过ISSN中国国家中心查询期刊的ISSN号进行辨别。作者可以通过网站<http://www.nlc.gov.cn>进入ISSN中国国家中心查询拟投期刊的ISSN号,当然,作者也可以通过拨打电话010-88548381,88545367,或者通过发邮件至issncn@nlc.gov.cn进行查询。

以上4种方法对作者而言可能会需要一定的辨识功夫,根据稿件编号直接查询则更简单。收到作者投稿后,如果进入的是虚假网站,会很快收到支付审稿费等费用通知,在支付费用前可向编辑部电话咨询,告知其稿件编号,学术期刊的编辑根据编号方式即刻就可判定其稿件是否投到虚假网站。

根据网址(URL)特征也可甄别网站的真假。URL由3部分组成,即资源类型、存放资源的主机域名、资源文件名。虚假网站一般有复杂的网络地址或单纯的网页结构。以黑龙江科技大学学报为例,其虚假网址有:

1) [2\) <https://www.zhazhi.com/qikan/gyjs/kygc/84409.html>;](https://se.warning.360.cn/warn/?DAUAAAFNAAAAAQAQBQTELQWrc5SmgYb4shdW/6iLeOzSBRJXhdr3glgqnfQu2+fY1Rgx1H7zUCaWey8GdklqbHCijKeW7FoB04eQZA6W/ZIIQ/QPg10DMsmECAJfdVBGYOZTcjzgj36jHsTQ0xl+fEsKGBAWcqIagC7i6P924kur1llyJalJfhcbmAjllmaCfwXYqMPR19emzCvNLRxrmGdVoKx8QJRDRk/6y3Kw7fOQIPM801ct/qGJ3JUUK+vLZqnOvtv/n7rXiKx0TxKtj8SNrqtbbSU8U2R8dX2q4irdGgU22BNzT1vInyBE/qZdLCKG6dvQMPnUAvpPNx5a5OFmtNn2gbmLolELBKlgzsh4UMQbrU5aiKBQd16zUrNKO2el9sNxatIgbLY40icKpUvITDud1RNJbscRdcEeVYbmdrXEL1rJFRhyvpm9gHqCFLHCZa;</p>
</div>
<div data-bbox=)

3) <http://www.qikanzj.com/kejiqikan/23069.html>;

4) <http://www.qikansky.com/b/kjiqk/19100.html>。

其中1)的网址存在博彩危险,网站2)~4)仅是单纯的html网页,这与正规网站的网址差别还是较大的。

2 期刊官网查询软件系统的设计

学术期刊内容是具有独创性的智力成果。技术进步为学术资源的传播利用带来极大便利^[9],网站、网刊、微信公众号是学术期刊自主传播的主要信息平台^[10],期刊出版者严守学术出版规范^[11],有着完善的学术不端审核机制^[12]。可是虚假网站的侵权,不同于

合法网站在内容传播上的知识产权冲突^[13],具有显著的主观恶意侵害。学术期刊虚假网站的运营是盗用学术期刊名义牟利,非法向作者征稿,并以审稿费或版面费的名义向作者收取费用问题,这种做法是严重的侵权行为。它不仅会误导作者,使作者蒙受经济损失,而且极大地损害被盗用期刊的名誉。同其他知识产权侵权纠纷一样,网站侵权证据的确定、赔偿数额的认定、侵权网站终端设备所在地以及案件的管辖也是审判实践中棘手的问题。作者和学术期刊出版者的权利主张不仅要面对实际法律问题、市场意识,还有技术门槛^[14]。采取技术保护或可是一种救急方式。

笔者根据国家新闻出版广电总局认定的2批学术期刊名录数据和新闻出版广电总局、中国知网等网站信息,通过网页爬虫技术,采用网页过滤算法,实现学术期刊网址查询。结合MFC和爬虫技术以期刊名、CN号、ISSN号等信息开发一种查询学术期刊投稿网址的系统。该系统由查询内容输入、查询类别选择和信息显示窗口组成。该软件具有可同时显示期刊的名称、CN号、ISSN号、出版语种、出版周期、投稿网址和期刊收录情况,尤其是准确高效地获取期刊的网址对于投稿作者至关重要。

采用通用性搜索引擎搜索期刊网址有遇到虚假网站的风险和系统不够智能的问题^[15-16],本设计可准确获取期刊官网信息及重点报道方向,通过访问网上的网页信息和相关链接,采用相关算法抓取期刊网址信息,软件系统总体的结构采取3层结构,即用户界面层、爬虫期刊官方网址的挖掘和数据库网址的更新。

2.1 用户层界面的实现 采用MFC Visual C++ 6.0开发工具通过用户界面的设计对文件操作。MFC类库可实现多线程调度与处理,通过CWinAPP类完成由线程多任务所引发的用户界面,其是CWinThread类的派生类,是用户接口线程的基本类,负责处理用户产生的事件和消息。CWinThread的对象用以维护特定线程的局部数据。CSpiderThread类的定义如下:

```
class CSpiderThread : public CWinThread
{
public:
    CSpiderThread(AFX_HINSTANCE pThreadParams);
    virtual ~CSpiderThread();
    PfnThreadProc, ThreadParams
    static UINT ThreadFunc(LPVOID lparam);
    BOOL ThreadRun(ThreadParams * vThreadParams);
    ...
};
```

软件的用户界面主要功能是实现作者查询期刊官

方网站及相关信息的需求,其基本数据结构如下:

```
class JournalInfo
{
public:
    JournalInfo();
    ~JournalInfo();
public:
    CString m_name;
    CString m_CN;
    CString m_ISSN;
    CString m_department;
    CString m_host;
    CString m_language;
    CString m_period;
    CString m_web;
    CString m_retrieval;
};
```

该数据结构用来记录基础的网址数据,包括期刊名称、CN号、ISSN号、主管单位、主办单位、出版语种、出版周期、投稿网址、期刊收录情况等信息。该数据结构同时也用来记录非法网址。当非法网址被检测到以后,可由该数据结构的对象记录,更新到相应的数据库中,在用户查询过程中,对于用户查询的网址,系统会首先在官网数据库和非法数据库中进行搜索,给出合理的结果。当没有用户搜索的数据信息时,再通过互联网搜索,并对搜索到的网址记录同时更新相应的数据库。

用户界面由以下步骤实现:

定义界面控制变量如下:

```
CEdit m_edit;
CButton m_radio_name;
CButton m_radio_cn;
CButton m_radio_issn;
CListCtrl m_list;
```

其中:变量m_edit允许用户输入查询关键词;3个按钮m_radio_name、m_radio_cn、m_radio_issn提供3个互斥的搜索方式,即按期刊名、CN号以及ISSN号搜索;m_list用来显示查询到的结果列表。

在虚函数OnInitDialog()中增加如下代码来初始化系统:

```
{
...
    m_list.InsertColumn(0, _T("类别"), LVCFMT_LEFT, 80);
    m_list.InsertColumn(1, _T("内容"), LVCFMT_
```

```
LEFT, 500);
m_radio_name. SetCheck( TRUE);
m_radio_cn. SetCheck( FALSE);
m_radio_issn. SetCheck( FALSE);
...
}。
```

这里设置了默认查询方式,即按期刊名称查询。用户输入关键词后,点击“查询”按钮,一旦用户查询成功,相应的结果会显示在 m_list 的列表中,详细信息包括期刊名称、CN 号、ISSN 号、投稿网址、期刊收录情况等主要信息。系统界面如图 1 所示。此外,系统允许用户将格式化的网址列表导入到数据库中。



图1 系统界面

2.2 期刊官方网址的挖掘 网页爬虫技术可以自动采集期刊官网信息,并且对已有的网页信息进行自动更新。该技术采用分布式和多线程技术,能够快速获取期刊网页,并建立索引。MFC 可以调用数据库函数,通过设计 API 程序实现与数据库管理器接口通讯。爬虫网址技术挖掘部分挖掘期刊 HTML 源文件,通过以下 4 步实现:一是判断 URL 是否具有虚假网址的特征,通过 HTTP 超文本传输协议抓取网页内容,创建 CInternetSession 对象;二是创建与 Internet 的会话和 CHttpConnection 对象,调用 CInternetSession 类成员函数 GetHttpConnection 来创建 CHttpConnection 对象;三是建立 HTTP 连接,建立 HTTP 连接之后可获得 CHttpFile 对象;四是抓取 HTML 源文件,通过 CHttpFile 类对符合查询条件的期刊网址的 HTML 源文件进行抓取、挖掘和分析。

HTML 基本结构如下:

```
<html >
<head >
<title >期刊官方名 </title >
</head >
<body >
```

期刊刊号、主管和主办等信息

```
</body >
</html >。
```

网页爬虫对 HTML 源文件的解析从网页的名称和显示部分进行。网页名称的查找是对源文件 <head > <title >期刊官方名 </title > </head > 标签部分的挖掘;网页的显示部分通过对源文件 <body > 与 </body > 标签中的内容进行查找、记录,完成 HTML 源文件的解析。网页爬虫定义的消息类型如下:

```
enumMSG_TYPE
{MSG_INIT_NIFORMTION_JOIN, //Spider 请求加入
MSG_SEVERFIND, //Spider 发现新网站
MSG_MY_IP_ADDRESS, //Spider 的 Ip 地址,发
MSG_PAGES_DOWNLOADADED, //Spider 搜索完成,发
}。
```

2.3 网址的更新与存储 利用数据库技术存储更新网址信息,通过 MFC 与网页爬虫技术获取期刊网站的网址、主管、主办、主要报道方向等信息,网页爬虫程序需要 MFC 相关协议获得相应的页面,解析期刊网页内容,分析其链接,以获取主管、主办等信息。当作者在查询框内输入相关关键词后,该系统将返回相关的期刊官方网站地址, MFC 通过导入 CApplication, CMyFont 等类将网址存储到 Excel 中。

2.4 软件的实现 在输入查询内容中输入期刊名“煤炭技术”,查询类别中选择“期刊名”,然后点击查询按钮,输出查询结果如图 2 所示。



图2 软件查询结果

由图 2 可见,它显示了“煤炭技术”期刊的名称、CN 号、ISSN 号、出版语种、出版周期、投稿网址和期刊收录情况。以“CN 号”和“ISSN 号”查询的方式与“期

刊名”的查询步骤类似。采用常规的搜索引擎搜索的结果如图3所示。



图3 搜索引擎的搜索结果

由图3可见,其真正的官方网站是第4个,对于没有经验的作者很难辨别其真假网站,而采用本文的软件避免了分辨的困难,在很大程度上避免了作者和编辑部受到侵害。

3 结束语

基于学术期刊 ICP 备案、搜索引擎“官网”认证、DOI 等信息,可以建立学术期刊官方网站的有效识别技术。可以通过向网站注册服务机构、ICP 备案机构、服务器所在机构、互联网违法和不良信息举报中心、网络警察举报等措施,对学术期刊假冒网站进行有效阻击。笔者所研制的学术期刊网址查询系统,使作者可以直接查找到学术期刊的真实网站,精准投稿,减少虚假网站的影响。该系统由《黑龙江科技大学学报》编辑部定期维护,力争数据采集更新的准确性。目前数据正在进一步完善中,待软件调试稳定,可从《黑龙江科技大学学报》网站首页下载使用。下一步将结合上述基本功能,扩展补充期刊的主要报道方向、特色及与相关数据库网站的链接,按作者的稿件关键词智能推荐期刊,以便更好地帮助读者投稿。本系统待进一步完善后可供广大作者、学术期刊和相关用户使用。

目前,学术期刊出版者在版权价值挖掘和网站运营能力方面尚待提升,在数字技术和网络平台环境下,

法律的滞后和已有规则的缺陷,抑制了学术期刊被虚假网站侵权后的维权积极性。切断学术期刊虚假网站的源头,通过技术创新手段,才能使版权产业得到有效保护,进而使知识创新的积极性得到鼓励。

4 参考文献

- [1] 刘永强,李园,马昕红,等. 学术期刊傻瓜式投稿模板设计方法[J]. 编辑学报, 2018, 30(2): 193
- [2] 王明华,沙勤,顾凯,等. 学术期刊独立网站网络传播环境特征及其品牌拓展[J]. 编辑学报, 2015, 27(2): 112
- [3] 孙昌朋,林萍,陈望忠. 百度竞价排名对中文核心期刊官网在百度搜索中的影响及思考[J]. 中国科技期刊研究, 2016, 27(8): 888
- [4] 黄锋,黄雅意,辛亮. 科技期刊假冒网站的防范和打击策略研究[J]. 中国科技期刊研究, 2016, 27(7): 740
- [5] 黄仲一,郭雨梅,朱雪莲. 科技期刊假冒网站问题分析及应对策略[J]. 编辑学报, 2016, 28(1): 50
- [6] 杨继涛,潘新社. 科技期刊对网络侵权象的防范措施[J]. 中国科技期刊研究, 2016, 27(10): 1078
- [7] 张颖. 科技期刊如何规避非法期刊网站的侵害[J]. 中国科技期刊研究, 2011, 22(3): 421
- [8] 贺嫁姿,温优华. 论文中介网站冒名期刊欺骗作者的危害与应对策略[J]. 中国科技期刊研究, 2016, 27(6): 608
- [9] 胡文杰,杨春明. 新形势下学术期刊产业联盟发展模式研究[J]. 编辑学报, 2018, 30(3): 227
- [10] 陈勇,郭伟. 媒体融合背景下科技期刊学术传播方阵的构建与探索[J]. 编辑学报, 2019, 31(2): 138
- [11] 田恬,陈广仁. 明确学术出版道德强化期刊编辑规范[J]. 编辑学报, 2017, 29(3): 206
- [12] 晁晓筠. 科技学术期刊在学术诚信建设中的作用[J]. 编辑学报, 2011, 23(4): 287
- [13] 岳玉荣. 网络环境下的知识产权冲突及技术保护[J]. 青海师范大学学报(哲学社会科学版), 2005(1): 16
- [14] 晁晓筠. 传统学术期刊数字化面临的著作权问题[J]. 编辑之友, 2002(4): 63
- [15] 康锋,唐志荣,陈丽琼,等. 研究生稿件作者自查表的编制与应用[J]. 编辑学报, 2019, 31(2): 192
- [16] 苏磊,蔡斐,袁睿. 新媒体环境中新型学术期刊网站设计实例分析[J]. 编辑学报, 2018, 30(6): 622

(2018-11-30 收稿;2019-07-01 修回)