

基于“文本-主题”双层次分析的网络虚假评论检测法研究^{*}

——以电子商务网站为例

赵衍^{1,2}

(1. 上海外国语大学信息技术中心, 上海 200083; 2. 上海外国语大学电子政务国际化研究中心, 上海 201620)

摘要: 首先从“重复度”和“主题相关度”两个维度对网络虚假评论进行了详细的分类; 基于此, 结合现有技术, 提出一种较为系统且成本较低的“文本-主题”双层次网络虚假评论检测方法; 将该方法应用于国内某电子商务网站的虚假评论检测, 获得了较理想的效果。

关键词: 网络虚假评论; 网络评论; 评论检测; “文本-主题”双层次分析

中图分类号: G35

DOI: 10.3772/j.issn.1673-2286.2015.09.009

1 引言

随着电子商务的深入发展, 电子商务网站成为网购的重要渠道。为了提升消费者购物体验, 提高网站用户的黏性, 电子商务网站都会提供用户评论 (Review) 功能, 让用户就购买过的某个产品发表使用感受, 并作为其他用户的购买参考。由于用户评论是表达用户使用感受的第一手资料, 具有很强的真实性和参考价值, 因此, 有70%左右的消费者相信网络评论^[1], 有67%以上的消费者在网购前会查看所要购买产品的评论^[2]。可以说, 电子商务网站的评论已经成为网购行为的主要参考信息。

商家也注意到网络评论对消费者的巨大影响, 因此, 一些商家开始试图操控网络评论, 从而影响消费者决策, 网络虚假评论 (Fake Review) 问题出现, 并日益严重。研究表明, 美国最大的B2C网站Amazon上有30%以上的产品评论是虚假评论; 美国最大的点评网站

Yelp上有16%的评论是虚假的^[3]; 我国的淘宝、京东等商家也承认其电子商务网站上存在大量的虚假评论。

网络虚假评论掩盖了产品的真实情况, 为购买者决策提供虚假信息, 扰乱市场信息传导机制的正常工作, 对社会的总体价值造成了损害^[4]。

2 文献综述

对网络虚假评论的研究最早开始于2000年, 但早期成果极少。此后, 从2007年起, 随着电子商务的蓬勃发展, 对网络虚假评论的研究呈现日益繁荣的趋势。从研究主体的角度来看, 对网络虚假评论的研究主要分为两大类。

2.1 基于文本技术的虚假评论识别

此类研究一般都将虚假评论识别视为二值分类

^{*} 本研究得到2013年上海市哲学社会科学规划课题 (编号: 2013ETQ001)、上海市教育委员会2014年科研创新项目 (编号: 14ZS070)、上海外国语大学“2013教学科研团队”项目、上海外国语大学“2014青年教师创新团队”项目 (编号: QJTD14ZY001)、上海外国语大学高层次人才发展计划 (编号: KX171260) 资助。

问题,从评论的文字和语义层面入手,分析网络虚假评论的用词、语句等字面特征,并对这些特征进行归纳,从而建构虚假评论的特征模型,并利用模型进行虚假评论识别。这一类研究主要使用监督学习的方法,先选择一定量的虚假评论样本进行标注和训练,得到虚假评论特征模型,再利用特征模型进行虚假评论识别。具体的识别算法有K近邻(K - Nearest Neighbor, KNN)分类算法、贝叶斯(Bayes)分类算法、支持向量机(Support Vector Machine, SVM)、Logistic回归(Logistic Regression, LR)算法等。

如Yoo^[5,6]等人从欺骗理论(Deception Theory)得到启发,从旅游网站上分别抽取了若干条真实评论和虚假评论,从语言结构(Language Structure)角度将真实评论和虚假评论进行特征对比,发现真实评论和虚假评论在词汇复杂度、第一人称使用频率、品牌提及频率及评论者情绪等方面存在不同,但研究者发现很难用结构化的方法去甄别真假评论。何海江^[7]等通过爬虫随机摘取网络评论、博客文章和BBS文档,分别选择其中一部分进行人工标注,建立训练集,采用Logistic识别垃圾评论,并将其与支持向量机(Support Vector Machine, SVM)的性能进行对比;Myle Ott^[8]等人利用亚马逊土耳其机器人(Amazon Mechanical Turk)获得了800篇评论,然后采用二元分类法,从文本风格识别、心理学欺骗检测等角度探测从字面上不容易被识别的虚假评论,经检验,准确度达到90%。Nitin Jindal和Bing Liu^[9-10]从Amazon网站上抓取了近600万条评论,并建立虚假评论样本进行训练,采用Shingle和Logic Linear Regression两种算法识别重复评论;Song Feng^[11]等通过基于概率的上下文无关文法规则特征分析,用SVM分类器对真实与虚假评论进行分类。

监督学习法虽然效果较好,但需要大量的人工标注,训练集的选择对结果影响很大,且成本较高,因此,有学者尝试用无监督学习法研究虚假评论。如谭文堂^[12]等将整个评论视为电阻网络,将虚假评论作为该网络上的语义离群点,设计一种基于电阻网络的虚假评论检测方法,该方法计算各节点(各条评论)对整个电阻网络(所有评论)的平均电能消耗的影响,建立“电离群因子”度量节点的离群度,从而识别主题无关评论,取得了一定的效果。但由于无监督学习方法在文本挖掘领域的研究尚不成熟,因此,这方面的研究文献较少。

也有学者试图综合监督学习和非监督学习的优

点,采用半监督学习的方法检测虚假评论。如郭利强^[13]在对博客文本进行统计分析的基础上,设计相关度、词组重复率、超链接数目、内容淫秽度、句子长度五个特征指标,给出一种半监督学习式网络垃圾评论检测方案和检测系统的框架,并进行实验验证;Fangtao Li^[14]等人则在构造虚假评论样本库的基础上,设计机器学习系统,对虚假评论样本进行识别,并利用Co-Training算法将学习结果应用于对其它评论的检测,取得了较好的效果。

2.2 基于虚假评论、虚假评论者、虚假评论对象等要素的虚假评论研究

网络虚假评论的具体情况非常复杂,仅通过字面分析的识别效果有较大的局限。因此,有学者将虚假评论、虚假评论者、虚假评论对象等要素结合起来进行分析,以提高虚假评论识别的效率和准确性。

如,C Dellarocas^[15]对用户的打分行为进行了聚类分析,发现信誉等级低的评论发表者所发表的虚假评论比例较高;王辉^[16]等人将观点作弊者分为个人作弊者和群体作弊者两类,分别研究了以评论为中心的作弊探测、以评论者为中心的作弊探测和以服务器为中心的作弊探测三种作弊方法;李弼程^[17]等人结合评论者名、评论发出的IP地址、发表评论时间及评论文本本身等四个特征量,提出了一种基于SHA-1技术的重复评论检测方法;Ee-Peng Lim^[18]等人对虚假评论者的特征进行归纳,并建立虚假评论者识别模型; Mukherjee和Liu B^[19-20]等人分析了虚假评论发布者的群体作弊行为,并开发了发现群体作弊行为的三个步骤;Jindal N^[21-22]等人深入分析了用户对产品的打分行为,从中发掘不合常理的用户打分模式;Wu Guangyu^[23]等人分析评论者发表的唯一的一条评论在某个产品的所有评论中所占的比例,以及这些评论在时间维度上的聚集程度,从而发现虚假评论者;李方涛^[24]从评论本身和评论作者进行虚假评论的识别,并验证从两个角度分析的Co-Training算法在虚假评论的识别上要优于传统的单角度半监督学习算法;Sihong Xie^[25]等重点研究单次虚假评论,从统计学角度考量了虚假评论和真实评论在“到达”时间上的差异,检测虚假评论可能发生的时间窗口;Theodoros Lappas^[26]将虚假评论视为一种攻击行为,分析了几种不同的攻击方式和不同方式下影响攻击效果的因素;孟美任和丁晟春^[27]将虚假商品评论发布者

的动机分为推销、诋毁、干扰和无意义四种,分析了每种动机的特点,并基于这些特点揭示了造假行为和隐藏行为;陈隆升^[28]等人利用支持向量机(Support Vector Machine)来提升操弄评论的分类绩效,并试图找出评论操弄行为的重要特征,选取了Samsung Galaxy Note II的用户评论验证了该方法的有效性;宋海霞^[29]等通过分析评论者的行为特征,提出一种自适应聚类虚假评论检测方法。

有学者将评论和评论对象结合,对角度对虚假评论进行了研究。如Nitin Jindal^[21]等分析了amazon.com上的用户评论,将用户评论的对象细化为产品、品牌、类别,将用户评论分为正面、负面和中性,对于用户的评论欺骗行为提出了3种假设,分别设计了这3种情况下识别虚假评论的方法。

也有学者将虚假评论、虚假评论者和评论对象三者进行综合分析。如李霄和丁晟春^[30]从评论、评论者和被评论的商品三个方面选择11个特征,使用支持向量机模型中4种常用的核函数进行垃圾商品评论的识别,对识别效果较好的RBF核函数中的参数C和 γ 进行优化,使得商品评论中的垃圾评论识别的准确率提高到78.16%,召回率提高到72.18%;Guan Wang^[31]等利用“Heterogeneous Graph”来分析评论者、评论和被评论的商店间之间的关系,通过对评论者可信度、评论的真实程度以及被评论商店的美誉度的综合量化分析,发现可能存在的虚假评论。

目前,对于网络虚假评论的研究日益增多,但是,从总体上来看,对于网络虚假评论的研究尚处于起步阶段。由于缺乏统一的研究框架和通用的概念模型,目前的研究缺乏系统性,研究成果比较零散,绝大部分研究成果的适用范围均局限于特定问题,不仅选取的样本量少,而且研究对象往往局限于某个特定领域的产品和服务的网络评论,特征抽取缺乏代表性,因此,研究成果缺乏通用性和系统性。

3 网络虚假评论分类

网络虚假评论既包括“网络水军”为了提升或降低某产品或某品牌的美誉度而作的正面或负面评论;也包括用户为了赚取经验值而作的无意义(即没有实质性内容,如随机字符等)评论;还包括利用电商平台进行广告宣传的广告贴等等。

对网络虚假评论的研究需要一个通用的、科学的

框架,能够处理大多数类型的网络虚假评论;同时,该框架也需具备较强的系统性,利用该框架,能够在网络虚假评论检验的“查全率”和“查准率”两个指标上有较好的表现。而要使得虚假评论检测系统既具“通用性”又具“系统性”,则首先必须对网络虚假评论进行科学、细致的分类,针对不同类型的虚假评论进行有区别的识别。

从评论内容本身来看,可以从多个角度对网络虚假评论进行细分:从文字层面的相似度,可以分为重复评论和非重复评论;从语义层面的相似度来看,可以分为相同语义评论和不同语义评论;从语义褒贬层面可分为赞扬性评论和诋毁性评论;从语义与主题相关性层面来看,可分为主题相关评论和主题无关评论。在实际情况中,这些分类之间又是相互交叉的,如有些评论是主题无关的赞扬性评论,有些评论是主题相关的诋毁性重复评论等等。如果以评论内容的重复与否为横坐标,以主题相关性与否为纵坐标,可用图1对网络评论进行简单分类。

由于评论只有存在明确的被评论对象才有实际意义和价值,因此,“真实”评论应该都是与评论对象主题相关的,与主题无关的评论基本上都可判定为虚

主题相关	真(可能性较大)	假(网络水军)
	假(无意义评论)	假(广告)
主题无关	内容不重复	内容重复

图1 网络虚假评论分类

假评论(即使评论者并不打算提供“虚假”评论,但由于该评论对浏览者并无参考价值,所以,依然可以归入“虚假评论”范畴)。其中主题无关的非重复评论一般可判断为无意义评论,大多是评论者为赚取积分随便敲入的字符;主题无关的重复评论一般是广告,评论者通过“复制-粘贴”的方式发布的。在后续的实证研究中也证明了此分析结果——主题无关重复评论一般都是类似“很好很好很好”、“可以可以”等无意义的形容词,甚至还有一些无意义字符,且评论中词语重复率高,评论普遍都很短;另外,超过97%的主题无关重复评论是广告。

主题相关的评论情况较为复杂。主题相关的非重复评论为真实评论的可能性较大,但不能就此肯定为

真, 还需作进一步分析和判断; 在大部分情况下, 主题相关的重复评论往往是网络水军发布的褒扬或贬低评论对象的虚假评论 (虽然也不排除少数情况下真实的评论者为了省事拷贝并粘贴了评论内容, 但在这种情况下, 也可以认为这些评论是虚假的, 因为这些评论并不能百分之百地表示评论者的想法)。因此, 从评论内容本身来看, 主题无关评论和内容重复评论都有极大的可能是虚假评论。

对于主题相关的非重复评论, 还需综合评论者、评论对象和评论的极性进行综合判断。如果某个或某群评论者对某个品牌有强烈的正面评价, 而对其它品牌都有一致的负面评价, 则基本可以判定这个 (些) 评论者是虚假评论者, 他 (们) 所发出的评论均可标记为虚假评论。

对于剩下的评论, 还需要再进一步分析其语义。由于虚假评论的发出者一般都没有与其所评论对象有过实际的接触 (如使用过所评论的产品、接受过所评论对象提供的服务等), 因此, 无法提供对评论对象的详细描述。同时, 这些虚假评论者为了显示其评论的真实性, 又往往会使用大量的第一人称和情感较为强烈的措辞。因此, 如果存在有如下情况的, 主题相关的非重复评论是虚假评论的可能性就比较大: 评论用词带有非常强烈的褒义或贬义; 评论中带有大量的第一人称, 如我、我们等; 评论只涉及产品、品牌或厂家, 而鲜有对评论对象属性的具体描述。目前, 学术界对于这方面的研究较多, 本文在此不作专门研究。

4 网络虚假评论识别流程

基于以上分析, 可以对网络虚假评论的识别过程进行梳理, 该识别过程主要涉及文本重复性检测、文本主题抽取与主题分析等技术。在开始识别前, 需做好以下准备工作:

(1) 通过网络爬虫, 摘取网络评论, 标注网络评论的属性, 主要包括评论发出者、评论的对象、评论极性 (好评、中评或差评) 和评论内容;

(2) 通过半自动的方法对评论对象设立主题词袋 (Bag of Words), 以方便主题识别。

然后, 先进行文本的重复性检测, 再进行主题相关性检测, 进行虚假评论的识别, 识别过程 (图2) 如下:

(1) 首先, 进行网络评论进行文本重复性检查 (语句重复), 如果为重复评论, 说明这些评论为拷贝或是

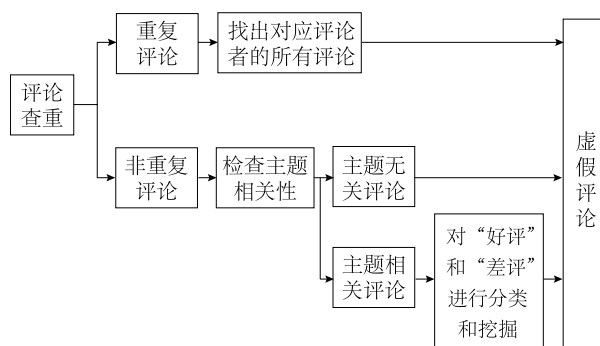


图2 虚假评论的识别过程

统一模板评论, 进而说明评论者的诚信存在问题, 找出这些评论者所发出的所有评论, 标记为虚假评论, 并标记这些评论者为虚假评论者; (虽然这样的规则会在一定程度上降低虚假评论检测的准确性, 但是可以提高召回率, 在后期结合一定的人工判断后依然可以保证其准确性);

(2) 对于非重复评论, 通过构建评论对象的主题“词袋”, 采用shingle算法, 判别评论与被评论对象之间的主题相关性, 如为主题无关评论, 说明评论没有参考价值, 将这些评论标记为虚假评论;

(3) 如为主题相关评论, 则结合评论者、被评论对象和评论极性进行聚类分析, 如一个 (类) 评论者总是对相同的评论对象发表情感相同 (同为正面或同为负面) 的评论, 则标记为虚假评论。

以上的识别结果会存在一定的误差, 如重复评论的发出者也许只发出过两条重复评论, 但他的所有评论都被标记为虚假评论。不过, 这种情况出现的几率是很小的 (因为在后续实验中, 发现重复评论者所发出的重复评论数均值在15左右)。而且, 信息检索的“查全”和“查准”两个指标本来就是负相关的。在海量的网络评论环境下, 牺牲一定的“查准率”来提高“召回率”是值得的。只要“查准率”的误差控制在可以容忍的范围内, 这样的一种判别方法是可以接受的。而且, 在后期加入适当的人工干预后, 可大幅度提高查准率。

5 实证分析

为了便于聚类和分析, 选取国内某B2C网站的手机评论作为数据源, 研究团队利用上海PN信息技术公司开发的“PN-i”智能网络爬虫程序, 爬取该网站上华为、中兴、联想和酷派4个品牌的智能手机的评论, 对每

个品牌，分别抓取2种评论数最高的型号的所有评论。
手机型号及其对应的评论数如下图所示（2014年9月22日14:50数据）：

表1 选取的研究对象及其评论列表

型号	总评论数	好评	中评	差评
酷派大神F1 (8297)	99467	95251	2746	1470
酷派8750	21229	18599	1679	951
华为荣耀6	19613	18729	554	330
华为荣耀3X pro	17857	16498	899	460
中兴U817	24823	22262	1730	831
中兴红牛 V5	14962	13480	914	568
联想黄金斗士S8	15917	14882	681	354
联想MA388	14507	13508	736	263
总计	228375	213209	9939	5227

为便于分析，只摘取评论中的所需信息，包括评论的内容（使用心得）、评论者名、评论者级别、评论极性，每条评论为一条记录，字段间以制表符为间隔标志，存放在review.txt文件中。

第一步：识别重复评论及评论者。

（1）首先进行文本重复性检测。目前，通过计算向量间的夹角的余弦公式在文本重复性检测中的使用已经较为成熟，因此，利用C++编程，在计算机上实现该算法。去掉评论中的所有标点符号，对评论内容进行重复性检测（文本内容超过95%的相似即判断为重复评论），并将内容重复的记录完整地取出，存放于文件copiedreview.txt中，共得到重复评论3736条。

（2）找出copiedreview.txt文件中所有重复评论的评论者，去重后，得到一个重复评论发出者的列表，存放于文件copiedreviewer.txt中，共有885人。

（3）找出review.txt文件中所有评论者存在于copiedreviewer.txt文件中的评论记录，标记为虚假评论，共得6682条评论。完整地取出这些评论，存放于文件copiedreviewerreview.txt中。

去除（1）和（3）中找到的虚假评论，review.txt文件中还剩217957条评论。

第二步：通过构造智能手机特征“词袋”，利用“shingle算法”，并利用C++编程，在计算机上实现该算法，识别主题无关评论。

通过查阅智能手机的设计资料，并结合具体使用

情况，共构建了与智能手机相关的19个主题，分别为屏幕、摄像头、外壳、电池、信号、速度、信号强度、无线网络、手感、美学、话筒、听筒、扬声器、闪光灯、按键、操作系统、质量、快递以及其它等，并通过每个主题构建“词袋”。主题及对应的“词袋”（部分）如表2所示。

表2 主题与词袋间的对应关系（部分）

主题	词袋
屏幕	屏幕/触摸/分辨率/成像/黑屏/字体/蓝屏/花屏/白屏/显示效果/细腻/模糊/清晰/触屏/看不清/耐磨/刮花……
摄像头	暗色效果/自拍/拍照/摄像头/清晰/清晰度/灵敏度/反应时间/像素/画面质量/光线暗/昏暗环境/照相/摄像/拍摄/全景/照相……
扬声器	喇叭/声音/噪音/杂音/破音/低音/高音/音质/放音乐/播放……
电池	电池/发热/续航/待机/寿命/安全性/容量/耗电太快/待机时间/电池/发热/发烫/可以用很久/没电/电量/毫安/锂电池……
其它	手机/价格/性价比/超值/包装/快递/不耍/功能/很好用/速度/贵/大品牌/结实/质感/性价比/质量/性能/反应/速度/外观/系统/边框/厚度/重量/有点儿重/不轻/工艺/灵敏/流畅/实用/物超所值/不错/很好/称手/满意/结实/高大上……

采用C++编写shingle算法，检测评论中语句落在“词袋”中的概率。为简单起见，如概率为零，则认为该评论为主题无关评论；否则就是主题相关。共找到主题无关评论49488条，取出后，存放于文件irrelevantreview.txt中。review.txt文件中剩余评论168469条。

第三步：评论分类与挖掘

考虑到有经济目的的虚假评论者只会给“好评”或者“差评”，而不会给出“中评”，因此，首先将review.txt中的“中评”（由于前面两步中已经去掉了“中评”1214条，故“中评”还剩8725条）去除，剩余评论159744条。

对负面评论和正面评论进行归类，分别存放于positivereview.txt和negativereview.txt文件中。考虑到在同一个电商平台上，虚假评论者的最终目的一定是宣扬某品牌，因此，他（们）一定给该品牌一致的正面评价；或者是在给该品牌一致的正面评价的同时，给其他品牌负面评价。因此，识别规则为：如某（类）评论者对其中一种品牌的评论均持一致的正面态度，或对某品

牌的评论均持一致的负面态度,则认为其是虚假评论者,然后找出该(这些)虚假评论者的所有评论,标记为虚假评论。

利用C++编写聚类程序,共识别虚假评论者683人,虚假评论7320条。取出这些评论存放于文件fakereview.txt中,review.txt文件中剩余评论152424条,再加上剩余的8725条“中评”,共剩“真实”评论161149条。

通过较为科学的分类,本研究分步骤地对网络虚假评论进行了层层过滤和识别,步骤如表3所示。

表3 智能手机虚假评论识别具体步骤

初始状态	共获得评论228375条,其中好评213209条,中评9939条,差评5227条
第一阶段	去除重复评论3736条和重复评论者发出的其它评论6682条,剩余评论217957条
第二阶段	去除主题无关评论49488条,剩余评论168469条
第三阶段	通过数据挖掘,去除虚假评论7320条,剩余评论“真实”161149条

经过三个阶段的识别后,共识别出虚假评论67226条,剩余(真实)评论161149,虚假评论率为29.44%。由于无法对228375条评论进行逐条的人工判断来计算查全率和查准率,因此,采用抽样法,从评论中随机抽取5组,每组1000条评论,进行人工分析,结果如下:

表4 查全率/查准率分析表

组别	样本数	人工检出虚假评论数	程序检出虚假评论数	程序检出的虚假评论中确为虚假评论数	查全率	查准率
1	1000	177	211	162	91.53%	76.78%
2	1000	278	285	251	90.29%	88.07%
3	1000	208	201	183	87.98%	91.04%
4	1000	301	311	283	94.02%	91.00%
5	1000	342	308	290	84.80%	94.16%
平均数					89.72%	88.21%

对随机抽样的查全率和查准率进行算术平均,得到本方法下的查全率为89.72%,查准率为88.21%,这样的数值在目前的虚假评论检测方法中处于中上等水平

(绝大部分方法的查全率和查准率在70%-96%之间)。但相比较而言,本方法的成本较低,使用较为简单,容易掌握,因此,具有很强的实践意义。

当然,本方法也存在一些不足。如,在对程序识别出的“虚假”评论和“真实”评论进行人工分析过程中,发现有一些评论存在模糊性,如,有评论为“合脚……”,按照本文的规则,为虚假评论,但按照语义来理解,此评论也可理解为“称手”,则就属于真实评论;再比如,有评论为“第二次购买了”,因为在“词袋”中没有此类主题词,因此,该评论看似与主题无关,但此评论可理解为评论者对该产品很满意,进行第二次购买,是与主题相关的正面评论;再如,评论“屏幕好,声音好,话筒好,好好好……”,虽然看似主题相关的真实评论,但在语义上却让人怀疑其真实性。

6 结语

本文通过分析虚假评论的产生机制,从而对虚假评论进行了较为系统的划分,并据此设计了一种分三个步骤的虚假评论识别方法。对比现有的有关研究,本研究所采用的方法虽较为简单,但更加科学和系统,识别率也较高。在识别的“成本”和“收益”间获得了较好的平衡。如果能够不断完善“词袋”,识别率将会得到进一步的提高。

但是,由于本研究只是提供一种简单实用的网络虚假评论检测方法,因此,在具体应用中还存在一些不精确之处,影响识别率。如,本研究所设计的虚假评论识别方法无法识别评论极性与真实评论内容相反的评论;也缺乏语义分析,难以识别不重复、不聚类的虚假评论。此外,对于一些在语义和主题相关性上存在模糊性的评论,本研究无法进行有效识别。因此,今后的研究可以再结合语义分析、情感分析等方法进一步提高识别精确度。

参考文献

- [1] Global Trust in Advertising and Brand Messages [EB/OL]. [2015-08-09]. <http://www.fi.nielsen.com/site/documents/NielsenTrustinAdvertisingGlobalReportApril2012.pdf>.
- [2] 中国互联网信息中心(CNNIC).中国网络购物市场研究报告[EB/R]. [2015-08-09]. <http://www.cnnic.net.cn>.
- [3] Luca M, Zervas G. Fake it till you make it: reputation, competition,

- and yelp review fraud[EB/OL].[2015-08-09] <http://people.bu.edu/zg/publications/fakereviews.pdf>.
- [4] Dellarocas C. Strategic manipulation of internet opinion forums: implications for consumers and firms [J]. *Management Science*, 2006, 52(10):1577-1593.
- [5] Yoo K H, Gretzel U. Comparison of deceptive and truthful travel reviews [C].//*Information and Communication Technologies in Tourism 2009*. Springer Vienna, 2009: 37-47.
- [6] Yoo K H, Lee Y, Gretzel U, et al. Trust in travel-related consumer generated media [C].//*Information and Communication Technologies in Tourism 2009*. Springer Vienna, 2009: 49-59.
- [7] 何海江,凌云. 由Logistic回归识别Web社区的垃圾评论[J].*计算机工程与应用*, 2009,45(23):140-143.
- [8] Ott M, Choi Y, Cardie C, et al. Finding deceptive opinion spam by any stretch of the imagination[C]. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, 2011: 309-319.
- [9] Jindal N, Liu B. Review spam detection[C].*Proceedings of the 16th International Conference on World Wide Web*, New York, USA: ACM, 2007: 1189-1190.
- [10] Jindal N, Liu B. Opinion spam and analysis[C].*Proceedings of the 2008 International Conference on Web Search and Data Mining*, New York, USA: ACM, 2008: 219-230.
- [11] Feng S, Banerjee R, Choi Y. Syntactic stylometry for deception detection[C]. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers*, Association for Computational Linguistics, 2012:171-175.
- [12] 谭文堂,朱洪,葛斌,等. 垃圾评论自动过滤方法[J].*国防科技大学学报*, 2012,34(5):153-157.
- [13] 郭利强. 网络博客空间中基于半监督学习的垃圾评论检测[J].*图书情报工作*, 2011,56(4): 52-55.
- [14] Li F, Huang M L, Yang Y, et al. Learning to identify review spam[C]. *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, Barcelona, Catalonia, Spain. 2011:2488-2493.
- [15] Dellarocas C. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior [C]. *Proceedings of the 2nd ACM Conference on Electronic Commerce*, New York, USA: ACM, 2000: 150-157.
- [16] 王辉,王晖昱,左万利. 观点挖掘综述[J].*计算机应用研究*, 2009,26(1):25-29.
- [17] 李真,林琛,李弼程. 基于Hash技术的重复性评论检测[J].*计算机应用*, 2009,29(S2): 263-266.
- [18] Lim E P, Nguyen V A, Jindal N, et al. Detecting product review spammers using rating behaviors[C]. *The 19th ACM International Conference on Information and Knowledge Management*, Toronto, Canada. New York, USA: ACM, 2010:939-948.
- [19] Mukherjee A, Liu B, Glance N. Spotting fake reviewer groups in consumer reviews[C]. *Proceedings of the 21st Annual Conference on World Wide Web*, Lyon, France. 2012:191-200.
- [20] Mukherjee A, Liu B, Wang J H, et al. Detecting group review spam[C]. *Proceedings of the 20th International Conference Companion on World Wide Web*, Hyderabad, India. New York, USA: ACM, 2011:93-94.
- [21] Jindal N, Liu B, Lim E P. Finding unusual review patterns using unexpected rules[C]. *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, Toronto, Canada. New York, USA: ACM, 2010:1549-1552.
- [22] Jindal N, Liu B, Lim E P. Finding a typical review patterns for detecting opinion spammers[R]. UIC, 2010.
- [23] WU G, Greene D, Smyth B, et al. Distortion as a validation criterion in the identification of suspicious reviews [C]. *Proceedings of the First Workshop on Social Media Analytics*, New York, USA: ACM, 2010:10-13.
- [24] 李方涛. 基于产品评论的情感分析研究[D]. 北京:清华大学, 2011.
- [25] Xie S H, Wang G, Lin S Y, et al. Review spam detection via temporal pattern discovery[C]. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, USA: ACM, 2012:823-831.
- [26] Theodoros Lappas. Fake reviews: the malicious perspective[C]. *17th International Conference on Applications of Natural Language to Information Systems*, Springer, 2012:23-34.
- [27] 孟美任,丁晟春. 虚假商品评论信息发布者行为动机分析[J].*情报科学*, 2013,31(10): 100-104.
- [28] 陈隆升,林瑞裕. 辨识科技产品操弄评论之研究[J/OL].[2015-08-09]. http://www.inf.cyut.edu.tw/AIT2013/ft_103.pdf.
- [29] 宋海霞,严馨,余正涛,等. 基于自适应聚类的虚假评论检测[J].*南京大学学报(自然科学)*, 2013,49(4):433-438.
- [30] 李霄,丁晟春. 垃圾商品评论信息的识别研究[J].*现代图书情报技术*, 2013(1):63-68.
- [31] Wang G, Xie S H, Liu B, et al. Yu. Identify online store review spammers via social review graph [J]. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2012, 3(4):1-61.

作者简介

赵衍, 男, 1978年生, 博士, 副教授, 研究方向: 文本挖掘、数据挖掘、知识管理等, E-mail: zy@shisu.edu.cn.

A Method for Detecting Web Fake Reviews Based on 'Text - theme' Double - Level Analysis:

A Case of Electronic Commerce Website

ZHAO Yan^{1,2}

(1.Information Technology Center, Shanghai International Studies University, Shanghai 200083, China)

(2.Center for Electronic Government Internationalization Research, Shanghai International Studies University, Shanghai 201620, China)

Abstract: In this paper, based on two dimensions - "Repetition" and "theme correlation", a detailed classification of fake web review is carried out firstly. Then, Combining with existing technology, a more systematic, but cheaper detection method - 'Text - theme' Double - Level Analysis Method - is presented. At last, this method is applied into the detecting of a Chinese e-commerce web site and got a good result.

Keywords: Web Review; Fake Web Review; Fake Review Detecting; 'Text - theme' Double - Level Analysis

(收稿日期: 2015-08-13; 编辑: 王立学)

■ 书 讯 ■

《科技报告体系构建研究》

为推进我国科技报告制度建设, 强化科技报告资源共享服务, 贺德方研究员率领中国科学技术信息研究所科技报告研究团队, 进行了国家社会科学基金重点项目“中国科技报告资源体系构建”(11ATQ006)研究, 并对20多年来中国科学技术信息研究所相关研究和实践进行了归纳、凝练、整理和补充, 最终形成了《科技报告体系构建研究》一书。

本书作为国家社会科学基金重点项目的研究成果, 总结了科技报告产生发展的管理历程、凝练了科技报告制度的建设路径、制订了科技报告资源的整合方案, 提出了科技报告体系的构建模式, 归纳了科技报告实践的操作过程。本书对各级科技计划管理人员强化科技计划项目过程管理具有借鉴作用, 对科研人员撰写高质量科技报告具有指导作用, 对各类科研机构做好科技报告呈交、推进科技项目的规范管理和机构知识库建设具有参考价值, 对图书信息机构做好科技报告深层次加工和收藏利用具有引导作用, 也可供高校信息管理、科技政策与管理等专业研究生学习参考。

《科技报告体系构建研究》于2014年12月由科学技术文献出版社出版, 定价78.00元。