垃圾商品评论检测研究综述

**张圣，伍星，邹东升**

**（重庆大学计算机学院，重庆 400044）**

**摘要:**商品评论对消费者的购买意愿有明显导向作用。欺诈者可杜撰评论来过度褒奖或恶意贬低商品，以此来促进己方或是打击对方的商品销售。垃圾商品评论检测成为了一项迫切需要的技术。本文首先将相关研究分为以评论内部（文本特征）为中心和以评论外部（文本特征）为中心两大类，然后分别综述它们在特征选择、学习方法上的研究进展，并整理了垃圾商品评论检测领域的常用评论数据集及其发展，在此基础上，展望了该领域的热点研究方向。

**关键词：**垃圾商品评论；文本特征；行为特征；评论数据集

Survey on review spam detection of product Reviews

**Zhang Sheng, Wu Xing,Zou Dong Sheng**

**（College of Computer Science,Chongqing University,Chongqing 400044, China）**

**Abstract:** The E-commerce website’s product reviews will have a guiding effect on the user's purchase intention.spammers may create fake reviews to artificially promote or demote target products.Therefore, techniques for detecting review spam has become an urgent need.This paper reviews the research progress of feature selection and machine learning method for the review centric（text features）and review’s external centric(behavior features) spam detection. And summarize common data sets in the field of opinion spam.At the end, some potential research direction are pointed out based on the review.

**Key words:** Review spam; Text Features; Behavior Features;Opinion Data set

1. **引言**

Web 2.0时代，用户在电商网站上发表了大量商品评论。这些评论不仅对该商品的潜在用户有明显导向作用[1]，还可以被商家用来改善产品、提升用户体验[2]。而受利益趋势，某些商家雇佣欺诈者发表含欺诈性质的评论：过度褒奖己方商品或者恶意贬低对手商品。

垃圾商品评论检测作为评论情感分析的预处理，主要特点是：可用于检测的特征选择面大、可用于检测的机器学习方法种类多、评论欺诈行为易于隐藏和改进。垃圾商品评论检测[3]从2007年被Jindal等提出至今，已经有大量研究从不同特征出发来解决该问题。研究路线可分为以评论内部为中心（关注评论文本本身的特征）和以评论外部为中心（关注评论者、评论团体和商品的特征）两类研究。垃圾商品评论检测研究难点在于缺乏权威数据集，而人工标注商品评论的难度很大。标注数据集的质量影响了传统监督学习算法和半监督学习算法在本领域的效果。需要解决的关键问题：1）绝大多数现有检测技术是在单个小数据集上完成的，而特征的有效性会随着数据集变更和数据集规模增大而改变，因此，模型对大数据集的兼容性是投入实际应用的关键。2）绝大多数现有检测技术是基于静态数据的，以离线方式执行各自的欺诈检测方案。欺诈行为很可能会在垃圾评论发布过程中明显暴露，而又被之后的评论事件掩埋，因此，实时检测也是急需解决的问题。

本文综述了垃圾商品评论检测的研究进展。第2节和第3节分别介绍了以评论内部和外部为中心的垃圾商品评论检测；第4节概括了本领域常用数据集；第5节总结全文，并展望了热点研究方向。

1. **以评论内部为中心**
   1. **以评论内部为中心的特征选择**

Table 1 review centric features

表1 以评论内部为中心的常用特征

|  |  |  |
| --- | --- | --- |
| 特征 | 文献 | 说明 |
| 词袋  特征 | [4],[5] | 通过n元语法记录一个词是否出现 |
| 词频  特征 | [6] ,[7] | 在词袋特征的基础上，还可记录一个词的出现次数 |
|  |  |  |
| 词法  句法  特征  词性  特征 | [8]  [11] | 词法特征用于统计偏好的词语类型、词语的平均长度等;句法特征用于获取评论者的句式写作风格  词性分析(POS--Part Of Speech)可挖掘文本的词性偏好 |
| 语义  特征 | [9] | 语义特征用于理解拥有近似语义的词语（wrong和error的相互替换不会影响到句子间的相似度） |
| 情感  特征 | [11], [34] | 通过统计正负向情感词数实现文本情感极性分析（LIWC--Linguistic Inquiry and Word Count） |
| 元  数据  特征 | [14] ,[15] | 包括评论长度、评论星级、评论反馈等 |

以评论内部为中心的研究从商品评论文本中选取特征（见表1）。Li[4]和Jindal[5]将一元语法和二元语法混合起来使用，把文本中连续的字和词语作为特征。Jindal[6]和Ott[7]在n元语法基础上统计了词频信息。Shojaee[8]引入了词法、句法特征，获取评论者的词语偏好和句式风格。Lau[9]引入语义特征来消除近义词相互替换对句子间相似度的影响。Guyon[10]评估了以上特征在欺诈检测中的重要程度，同时研究了特征选取数对结果产生的影响，发现移除一些贡献较小特征可避免过拟合。Ott[11]加入了心理学语言特征（LIWC），通过统计因果词、情感极性词等在全文中所占百分比来为评论检测提供文本情感倾向特征，它比较了词性分析（POS）、情感极性分析（LIWC）和二元语法这3组特征的性能，发现LIWC结合二元语法在支持向量机（SVM-- Support Vector Machine）中可以达到最佳检测效果。[Banerjee](http://xueshu.baidu.com/s?wd=author%3A%28Snehasish%20Banerjee%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)[12]发现豪华、便宜和中等这三档酒店中，真实和虚假评论在文本的复杂度（词法句法特征）和夸张表达（情感特征）等特征上的差异非常明显,细分领域能有效提高检测精度。Snehasish[13]从评论的情感特征入手，发现积极评论、消极评论和中性评论在垃圾商品评论检测中的差异性主要体现在可理解性和特异性特征。Li[14]和Hooi[15]主要从评论星级、评论反馈、评级分布等元数据特征来度量评论的可信度。

**2.2 以评论内部为中心的学习方法**

研究者将文本特征和元数据特征用于机器学习算法，根据研究对数据集里标注数据的需求程度不同，将评论内部为中心的学习方法分为以下三类。见表2:

Table 2 review centric machine learning method

表2 评论内部为中心的学习方法分类

|  |  |  |
| --- | --- | --- |
| 类型 | 文献 | 说明 |
| 监督  学习 | [3],[8],[11],  [16],[12],[13],[17] | 需要大量标注数据 |
| 无监督  学习 | [9] | 仅需要无标注数据 |
| 半监督  学习 | [4],[41],[18],[20],[19] | 需要少量标注数据和大量无标注数据 |

垃圾商品评论检测能用监督学习算法来解决，它可被看成二分类问题:真实评论和虚假评论。Jindal[3]和Lin[16]通过w-shingling算法计算两个文档的相似度，将相似度极高的评论作为正例（虚假评论），剩余评论作为负例，通过逻辑回归对未标注评论进行分类。Lin[16]在Jindal[3]的基础上，将评论相似度分成三类：评论与该评论者其它评论的相似度、评论与同商品其它评论的相似度、评论与该商品以外评论的相似度。Shojaee[8]用序列最小优化算法模型改进了用于欺诈检测领域SVM分类器算法性能（目标函数最优化问题）。吴[17]根据中文商品评论的特点提取了正负向情感特征、词频特征（品牌词个数）等9个特征来构建垃圾商品评论检测模型，通过关联规则的评论名词模式来提高主题词（评论中与主题相关度最高的名词）的识别精度，最后通过逻辑回归来完成分类算法。

Lau[9]为了克服垃圾商品评论检测领域标注数据匮乏的瓶颈，将无监督学习和基于语义特征的语言模型结合起来，通过关联规则挖掘来提取上下文相关的关联知识，核心思想是通过概率语言模型挖掘出语义上相似度很高的评论来作为重点检测对象，再结合SVM算法完成分类。这种方法很适合检测垃圾评论中的复制评论。不过，目前无监督学习的检测精度还不够高，且难以解决新出现的虚假评论与已标注的非虚假评论相似度高的问题。

针对无监督学习检测精度不高和监督学习数据匮乏的缺点，半监督学习算法被提出。Li[4]将2-视图的协同训练模型co-training运用到垃圾商品评论检测领域。用2组完全不同的特征（一组以评论内部为中心，一组以评论外部为中心）分别训练2个算法相同的分类器，co-training过程中，只有两个分类器同时将未标注数据集中某条数据通过为正例或负例才算标注生效。分类环节中，第三个分类器（包含前两个分类器的所有特征）使用前两个分类器准备的样本数据作为训练集，训练分类器并完成检测。PU-Learning算法是另一种用于垃圾商品评论检测的半监督学习模型,它最明显的优势是在训练集中不包含负例的情形下，所训练的分类器仍然可以获得较好的性能。Li[18]和Hernández[19]根据已标注正样本P和未标注样本U找出可靠负样本集合，再利用正负样本迭代训练得到二分类器。PU-Learning执行过程中，未标注数据集里存在一部分容易被错误标注的数据，[任](http://xueshu.baidu.com/s?wd=authoruri%3A%28a8d289bad7bbfbec%29%20author%3A%28%E4%BB%BB%E4%BA%9A%E5%B3%B0%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited)[20]为了优化PU-Learning对这类样例的聚类效果，首先基于狄利克雷过程混合模型对其进行聚类，再用多核学习算法将特征映射到了高维空间来提高区分度。

1. **以评论外部为中心**
   1. **评论外部为中心的特征选择**

Table 3 review’s external centric features

表3以评论文本外部为中心的常用特征

|  |  |  |
| --- | --- | --- |
| 特征 | 文献 | 说明 |
| 时间特征 | [43],[22],[15]  [42],[33] | 欺诈者的评论数通常有短期内爆发的倾向，欺诈团伙倾向于短时间内在同一个商品集合下发布评论 |
| 位置特征 | [14] | 测试同一评论者前后两条评论IP位置的变化速度，若速度快得不合常理，则该评论者很可能是欺诈者 |
| 评级偏差特征 | [15] ,[26]  [27] | 由于欺诈者有歪曲、误导评论星级分布的倾向，欺诈者与平均评级间的偏差比可信评论者要大 |
| 团伙特征  品牌特征 | [31],[9],[30]  [32],[23],[24]  [3] ,[5] | 欺诈团伙会串通起来共同塑造行为，检测的常用方法是频繁项集挖掘  若评论者发布了与商品所属品牌关联不大的评论，或者在不同品牌和商品中使用了相似度很高的评论，则很可能是欺诈者 |

Mukherjee[21]发现通过评论外部特征来寻找虚假评论要优于仅根据评论内部特征寻找虚假评论，以评论外部为中心的常用特征见表3。

Jindal[6]发掘了一些可疑评论的图模型和欺诈者的行为模式（评级偏差特征），经验证这些特征与欺诈活动有非常高的关联性。Fei[22]发现了一个用于区分欺诈者与正常评论者的关键特征——突发评论（Bursty Reviews）:在很短的时间里，某个特定的商品或是某个特定品牌下的商品集合呈现出评论数爆发增长的情况。Hooi[15]在突发评论基础上做出改进，引入了垃圾商品评论的时间特征和评级偏差特征。而Li[7,26]以评论发布时间和频率为主要线索，发现欺诈者和非欺诈者的评论虽然都服从双峰分布，但是在时间轴上的分布有明显区别。随着商用垃圾评论过滤器的流行和真实评论数据量的增长，研究者开始与大型电商网站合作，获取用于垃圾商品评论检测实验的真实数据。研究者发现真实数据集中的时间序列和空间位置信息可以作为检测欺诈者的关键特征。Li[14]通过大众点评网的数据，挖掘出一些垃圾商品评论检测的模式：1）来源特征：从PC端发出的垃圾商品评论占大多数。2）时间特征：周二至周四发出的垃圾商品评论较多。3）位置特征：所有的餐馆评论都采集自上海，IP离评论地（上海）越远，越有可能是垃圾商品评论。4）平均旅行速度：通过IP变化测试用户移动速度，不合理的很可能是垃圾商品评论。Mukherjee[23]通过挖掘团伙的评级偏差、文档相似性等频繁模式来寻找潜在的欺诈团伙。Lin[24]和Hernández[19]将重复评论和近似评论作为欺诈训练样本，并加入了以下特征：一个评论者在某一商品下的评论频率、评论者评论频率和一个商品的被评论频率。Jindal[3]选取以下特征来剔除垃圾商品评论：1）评论长度。2）品牌与商品的对应性。3）评论与商品的联系紧密程度。而Jindal[5]在前文的基础上，将所选特征扩充到了36个，其中包含评论文本特征、评论人特征和商品特征。

**3.2 评论外部为中心的学习方法**

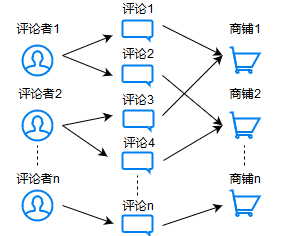
Mukherjee[25]将Yelp网站垃圾评论过滤器在真实环境下的检测效果与基于语法特征的模型做对比，证明了基于评论外部特征的模型优于基于语法特征的模型。根据评论外部特征所属实体（评论者、评论团伙和商品）将本节内容分成三类。见表4：

Table 4 review’s external centric machine learning method

表4 评论文本外部为中心的学习方法分类

|  |  |  |
| --- | --- | --- |
| 类型 | 文献 | 说明 |
| 评论者  实体 | [6],[22]  [21][26] | 欺诈者与可信用户在行为特征上有较为显著的区别 |
| 评论团  伙实体 | [31],[30]  [23],[24] | 欺诈团伙会共同塑造行为，以免出现可疑点.通常用频繁项集挖掘出可疑的评论团体，再将这个团体中所有成员在行为特征上的共性做为最核心的问题来考虑 |
| 商品实体 | [3],[5]  [33] | 将品牌与商品的联系、商品被评论的数量和频率等特征做为考量的重点 |

Mukherjee[21]通过评论图（见图1）模型来描述评论者、评论和商铺之间的关系，找出欺诈评论者和欺诈评论。具体来讲，评论图模型由三类节点组成：评论节点、评论者节点和商铺节点，将各节点的可信度量化为分值形式并捕捉节点之间的关系。再根据如下步骤来计算可信度分值：1）人为初始化商铺信用分和评论者信用分。2)计算每个评论的可信分。3)根据每一评论的可信分，反过来计算评论者的信用分。4）根据商铺所拥有的评论者可信分和评级偏差计算商铺可信分。5）用新的商铺信用分和评论者信用分去更新每个评论的可信分。6）重复操作直到迭代结束。若计算出的节点分值高于阈值，则将该节点判定为可信节点,反之亦然。Wang[26]用评论图模型迭代的来计算各节点的可信度分值，Wang[27]在评论者、评论和商家节点中，增加了更多量化指标来提升可信分的计算精度。



**Figure 1 relationships and nodes of graph model**

**图1.图模型的节点和关系[28-29]**

Fei[22]以突发评论为基础，结合置信度传播算法（Loopy Belief Propagation）和马尔可夫随机场（Markov Random Field）提出了基于突发评论的垃圾商品评论检测模型。Hooi[15]提出了基于用户评级行为（包括了评级频率和评级分布）和突发评论的贝叶斯推理，以此来度量用户的可疑程度。Santosh[28]分析了时间序列与垃圾商品评论的联系，通过Yelp评论数据发现了多个时间序列模式与垃圾商品评论发布频率的联系。采用向量自回归模型来预测不同的欺诈策略（早期欺诈、中期欺诈和晚期欺诈等）下的垃圾商品评论分布频率，将推导出的时间序列特征用于Yelp过滤器，提高它欺骗性评级的功能，该框架进一步的验证了Yelp过滤器的合理性，也在时间序列上改良了垃圾商品评论检测的效果。此外，Najada[29]发现在大量的在线评论中，虚假评论通常比真实评论要少很多，这就导致了高度的数据分布不平衡性，数据分布不平衡会影响分类器的性能，因此，在训练模型时很有必要考虑这一因素，可以对虚假评论采用高频率随机采样、真实评论采用低频率随机采样的方法来克服数据分布不平衡。

欺诈者往往都是以团伙的形式协同完成任务，在多个商品下发布垃圾商品评论，这种串通起来的欺诈行为是有迹可循的，团伙欺诈检测将欺诈团伙成员的共性作为核心问题来考虑。Mukherjee[23]通过挖掘到的团伙欺诈者特征来计算欺诈影响因子，并用SVM排序算法完成垃圾商品评论检测。Xu[30]提出了基于联合概率的潜在团伙欺诈统计模型，进一步挖掘团伙欺诈的共性。Mukherjee[31]为了更全面的描述团队、个人与商品的联系，提高挖掘团伙欺诈的精度，从三个二元关系(团队与个人，团队与商品，个人与商品)着手建立模型，再用频繁项集挖掘来找出潜在欺诈团伙。Lim[32]通过逻辑回归模型来实现评级行为，用以评价用户的欺诈嫌疑程度。

商品实体为中心的检测将品牌与商品的联系、商品被评论的数量和频率等特征做为考量的重点。Heydari[33]构建了一个包含以上特征的时间序列模型，通过时间窗口滑动来捕获每个商品在时间轴上的评论可疑模式。

1. **常用数据集**

Table 5 commonly used data set

表5 欺诈检测领域常用数据集

|  |  |  |
| --- | --- | --- |
| 数据集 | 文献 | 说明 |
| Amazon  商品评论数据集 | [3] ,[33]  [26] ,[30] | 伊利诺斯大学芝加哥分校(UIC—University of Illinois at Chicago)的Liu等收集了2006年开始的部分Amazon商品评论数据并提供了API接口，主要包括评论信息、商品信息和评论者信息 |
| 黄金数  据集  大众点评数据集 | [34] ,[38]  [39]  [43]，[14]  [42] | Myle Ott等采集了TripAdvisor上最热门的20家旅馆的评论,从中挑选可信评论，并将这些数据与Amazon Mechanical Turk(AMT)构造的虚假评论结合起来，构造了一个平衡的数据集  国内的大众点评与bing liu团队联合，开展真实大数据集下的中文欺诈评论检测 |
| 大型电商网站数据集 | [15] , [28]  [37] | 国外电商网站，如eBay,IMDB,flipkart,resellerratings等网站上的评论都被用于垃圾商品评论检测领域 |

在垃圾商品评论检测领域发展早期，研究大都采用Bing Liu等整理的Amazon 商品评论数据集，主要原因在于它涵盖了评论信息、评论者信息和商铺信息，且评论数据量大，种类多样。此后，Ott[34]构造了一个含标注且正负例平衡的数据集，被称为“黄金数据集”。但Mukherjee[25]通过实验证明，Ott所构造数据集训练出的模型在实验数据环境下检测效果较好，但是在真实数据环境下检测效果很差，这是由于文本特征在两种环境下的差异性造成的。因此，数据集的获取逐渐转向真实环境，研究者开始倾向于与大型电子商务网站合作（大众点评、eBay、flipkart等），这样做不仅可以使研究者获取大规模真实环境下的评论数据，也能够帮助大公司改进垃圾评论过滤器，提高网站的评论质量。

考虑到缺少权威标注数据集是欺诈检测领域发展过程中的一个巨大阻碍，且能与大型电商网站合作的研究者较少。除了以上这些主流的数据集，Shojaee[35]提出了一个框架，可以根据评论人信息、评论产品信息以及元数据，自动完成在线标注评论数据。

此外，使用已有工具来实时更新语料库、将垃圾商品评论检测向实时检测的方向推进也是数据集上的重要发展方向。SAMOA已经被用于实时分析Twitter流数据[36]，评论被源源不断的添加进语料库，让机器学习模型随着新数据流的加入而变化。

1. **总结与展望**

电商网站上的商品评论对于商家和消费者都十分重要。由于利益的趋势，商家开始雇佣欺诈者发表欺诈性质的评论，过度褒奖己方的商品或者恶意诋毁竞争对手的商品。本文综述了以评论内部和外部为中心的垃圾商品评论检测以及常用的数据集。

从目前的发展趋势看，垃圾商品评论检测今后的工作可围绕以下几点展开:

1）实时性:Ye[37]提出了一种基于时间序列的多元指示信号检测模型，它可以随着时间推移保持数据集的实时更新，动态的挑选最能区分虚假评论与真实评论的一组特征作为指示信号并实时监控，实现了实时的垃圾商品评论检测。

2）深度学习:更充分的利用数据集来学习模型，不再需要对数据集进行特征工程。 Ren[38]提出了一个卷积神经网络模型，它能从文档级检测含欺骗性的评论。神经网络模型相比传统的离散模型,具有更强的泛化能力。Li[39]在文档级神经网络模型的基础上尝试了多组特征组合，通过参数调优来提升模型的性能。

3）跨站点与跨语料: Yuan[40]挖掘多个宏观水平时间的情感模式和多个网站的微观层次特征。设计出跨站点情感时间序列来捕捉可疑评论攻击，实现跨站点的检测垃圾商品评论。进一步的，跨中英文语料库的垃圾商品评论检测也是一个颇具潜力的发展方向。

商品评论对商家和消费者都有非常重要的应用价值。然而在真实环境中,垃圾商品评论的普遍性、伪装性和团体性对已有检测技术提出了新的挑战,有很多新的方法技术需要进一步的探索和研究。

**参考文献:**

2. Samha A, Li Y, Zhang J. Aspect-based opinion extraction from customer reviews[J]. International Journal of Computer Science & Information Technology,2014, 6(3).
3. [Isidro Peñalver-Martinez](http://www.sciencedirect.com/science/article/pii/S0957417414001511), [Francisco Garcia-Sanchez](http://www.sciencedirect.com/science/article/pii/S0957417414001511), [Rafael Valencia-Garcia](http://www.sciencedirect.com/science/article/pii/S0957417414001511).Feature-based opinion mining through ontologies[J]. Expert Systems with Applications, 2014,[Volume 41, Issue 13](http://www.sciencedirect.com/science/journal/09574174/41/13), Pages 5995–6008.
4. Jindal N, Liu B. Review spam detection[C]. Proceedings of the 16th international conference on World Wide Web,2007,ACM.
5. Li F, Huang M, Yang Y, Zhu X.Learning to identify review spam[C]. IJCAI Proceedings-International Joint Conference on Artificial Intelligence,2011,vol 22, No. 3., p 2488.
6. Jindal N, Liu B.Opinion spam and analysis[C].Proceedings of the 2008 International Conference on Web Search and Data Mining,2008, (pp. 219–230).
7. Jindal N, Liu B, Lim EP.Finding unusual review patterns using unexpected rules[C]. Proceedings of the 19th ACM international conference on Information and knowledge management,2010,(pp. 1549–1552).
8. Ott M, Cardie C, Hancock JT.Negative Deceptive Opinion Spam[C]. HLT-NAACL,2013, pp 497–501.
9. Shojaee S, Murad MAA, Bin Azman A, Sharef NM, Nadali S.Detecting deceptive reviews using lexical and syntactic features[C]. Intelligent Systems Design and Applications (ISDA), 2013, 13th International Conference on (pp. 53–58).
10. Lau RY, Liao SY, Kwok RCW, Xu K, Xia Y, Li Y.Textmining and probabilistic language modeling for online review spam detecting[J]. ACM Trans Manage Inf Syst,2011, 2(4):1–30.
11. Guyon I, Elisseeff A.An introduction to variable and feature selection[J]. JMLR.org, 2003, 3(6):1157-1182.
12. Ott M, Choi Y, Cardie C, Hancock JT.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-Volume 1,2013, (pp. 309–319).
13. [S Banerjee](http://xueshu.baidu.com/s?wd=author%3A%28Snehasish%20Banerjee%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [AYK Chua](http://xueshu.baidu.com/s?wd=author%3A%28Alton%20Y.K.%20Chua%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels[J]. Journal of Information Science, 2016.

1. [Snehasish Banerjee](http://www.emeraldinsight.com/author/Banerjee%2C+Snehasish), [Alton Y.K. Chua](http://www.emeraldinsight.com/author/Chua%2C+Alton+YK).Theorizing the textual differences between authentic and fictitious reviews: Validation across positive, negative and moderate polarities[J]. Internet Research,2017,Vol. 27 Issue: 2, pp.321-337.
2. Li H., Chen Z., Mukherjee A., Liu B, Shao J.Analyzing and Detecting Opinion Spam on a Large-scale Dataset via Temporal and Spatial Patterns[C]. In: Proceedings of Ninth International AAAI Conference on Web and Social Media,2015.
3. B. Hooi, N. Shah, A. Beutel, S. Gunneman, L. Akoglu, M. Kumar, D. Makhija, C. Faloutsos.Birdnest: Bayesian inference for ratings-fraud detection[J]. arXiv preprint arXiv:1511.06030, 2015.
4. Lin Y., Zhu T., Wang X., Zhang J, Zhou A.Towards online anti-opinion spam: Spotting fake reviews from the review sequence[C]. 2014 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM) ,2014,(pp. 261–264).
5. Wu Hanqian,Zhu Yunjie,Xie jue.[Detection](http://www.cnki.com.cn/Article/CJFDTOTAL-DNDX201503004.htm) model of effectiveness of Chinese online reviews based on logistic regression[J].Journal of Southeast University(Natural Science Edition),2015,l(3).
6. Li H, Liu B, Mukherjee A, Shao J.Spotting fake reviews using positive-unlabeled learning[J]. Computación y Sistemas, 2014,18(3).
7. Hernández D, Guzmán R, Móntes y Gomez M, Rosso P.Using PU-learning to detect deceptive opinion spam[C].Proc. of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis.2013, (pp 38–45).

1. [Ren](http://xueshu.baidu.com/s?wd=authoruri%3A%28a8d289bad7bbfbec%29%20author%3A%28%E4%BB%BB%E4%BA%9A%E5%B3%B0%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited" \t "_blank) Yafeng,[Ji](http://xueshu.baidu.com/usercenter/data/author?cmd=authoruri&wd=authoruri%3A%28a009fb14d7cf2266%29%20author%3A%28%E5%A7%AC%E4%B8%9C%E9%B8%BF%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2) Donghong,[Zhang](http://xueshu.baidu.com/s?wd=authoruri%3A%288a0ca3b0d6b80b17%29%20author%3A%28%E5%BC%A0%E7%BA%A2%E6%96%8C%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited) Hongbin,[Yin](http://xueshu.baidu.com/s?wd=authoruri%3A%28660fb7648f52b237%29%20author%3A%28%E5%B0%B9%E5%85%B0%29%20%E8%B4%B5%E5%B7%9E%E5%B8%88%E8%8C%83%E5%A4%A7%E5%AD%A6%E6%95%B0%E5%AD%A6%E4%B8%8E%E8%AE%A1%E7%AE%97%E6%9C%BA%E7%A7%91%E5%AD%A6%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited) Lan.Deceptive Reviews Detection Based on Positive and Unlabeled Learning[J].Journal of Computer Research and Development,2015,l(3).
2. Mukherjee A, Kumar A, Liu B, Wang J, Hsu M, Castellanos M, Ghosh R.Spotting opinion spammers using behavioral footprints[C]. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining,2013, (pp. 632–640).
3. Fei G, Mukherjee A, Liu B, Hsu M, Castellanos M, Ghosh R.Exploiting Burstiness in reviews for review spammer detection[C]. ICWSM,2013:175–184.
4. Mukherjee A, Liu B, Wang J, Glance N, Jindal N. Detecting group review spam[C]. Proceedings of the 20th international conference companion on World Wide Web.2011.
5. Lin Y, Zhu T, Wang X, Zhang J, Zhou A. Towards online review spam detection[C].Proceedings of the companion publication of the 23rd international conference on World Wide Web companion,2014,(pp. 341–342).
6. Mukherjee A, Venkataraman V, Liu B, Glance N. What yelp fake review filter might be doing[C].Seventh international AAAI conference on weblogs and social media,2013.
7. Wang G, Xie S, Liu B, Yu P. Review graph based online store review spammer detection[C]. IEEE 11th international conference on data mining (ICDM),2011.
8. Wang G, Xie S, Liu B, Yu PS.Identify online store review spammers via social review graph[J]. ACM Transactions Intelligent Systems & Technology,2012,3(4):61.
9. K. Santosh, A. Mukherjee.On the temporal dynamics of opinion spamming - case studies on yelp[C]. In: WWW, 2016.
10. Al Najada H, Zhu X.iSRD: Spam review detection with imbalanced data distributions[C]. Information Reuse and Integration (IRI), 2014 IEEE 15th International Conference on (pp. 553–560).
11. [C Xu](http://xueshu.baidu.com/s?wd=author%3A%28Xu%2C%20Chang%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [J Zhang](http://xueshu.baidu.com/s?wd=author%3A%28Zhang%2C%20Jie%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).Towards collusive fraud detection in online reviews[C]. [IEEE International Conference on Data Mining](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=confuri%3A%287873c0716cc4f680%29%20IEEE%20International%20Conference%20on%20Data%20Mining&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited), 2015:1051-1056.
12. Mukherjee A, Liu B, Glance N.Spotting fake reviewer groups in consumer reviews[C]. Proceedings of the 21st international conference onWorld Wide Web,2012, (pp. 191–200).
13. Lim P, Nguyen V, Jindal N, Liu B, Lauw H.Detecting product review spammers using rating behaviors[C].Proceedings of the 19th ACM international conference on Information and knowledge management.2010.
14. Heydari A, Tavakoli M, Salim N.Detection of fake opinions using time series[J].Expert Systems with Applications.2016; 58:83-92.
15. Li J, Ott M, Cardie C, Hovy E.Towards a general rule for identifying deceptive opinion spam[C]. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics,2014, pages 1566–1576.

1. [S Shojaee](http://xueshu.baidu.com/s?wd=author%3A%28Somayeh%20Shojaee%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson" \t "_blank)，[A Azman](http://xueshu.baidu.com/s?wd=author%3A%28Azreen%20Azman%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[M Murad](http://xueshu.baidu.com/s?wd=author%3A%28Masrah%20Murad%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[N Sharef](http://xueshu.baidu.com/s?wd=author%3A%28Nurfadhlina%20Sharef%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[N Sulaiman](http://xueshu.baidu.com/s?wd=author%3A%28Nasir%20Sulaiman%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).A Framework for Fake Review Annotation[C]. 17th UKSIM-AMSS IEEE International Conference on Modelling and Simulation,2015.
2. [N Kourtellis](http://xueshu.baidu.com/s?wd=author%3A%28Nicolas%20Kourtellis%29%20Yahoo%20Labs&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [GDF Morales](http://xueshu.baidu.com/s?wd=author%3A%28Gianmarco%20De%20Francisci%20Morales%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [F Bonchi](http://xueshu.baidu.com/s?wd=author%3A%28Francesco%20Bonchi%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)Scalable.Scalable online betweenness centrality in evolving graphs[J]. [IEEE Transactions on Knowledge & Data Engineering](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=journaluri%3A%28b49dd15379f5ffb2%29%20%E3%80%8AIEEE%20Transactions%20on%20Knowledge%20%26%20Data%20Engineering%E3%80%8B&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited), 2015, 27(9):2494-2506.
3. [Ye, Junting](http://adsabs.harvard.edu/cgi-bin/author_form?author=Ye,+J&fullauthor=Ye,%20Junting&charset=UTF-8&db_key=PRE), [Kumar, Santhosh](http://adsabs.harvard.edu/cgi-bin/author_form?author=Kumar,+S&fullauthor=Kumar,%20Santhosh&charset=UTF-8&db_key=PRE), [Akoglu, Leman](http://adsabs.harvard.edu/cgi-bin/author_form?author=Akoglu,+L&fullauthor=Akoglu,%20Leman&charset=UTF-8&db_key=PRE).Temporal opinion spam detection by multivariate indicative signals[J]. arXiv preprint [arXiv: 1603.01929](https://arxiv.org/abs/1611.06625),2016.
4. [Y Ren](http://xueshu.baidu.com/s?wd=author%3A%28Yafeng%20Ren%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [D Ji](http://xueshu.baidu.com/s?wd=author%3A%28Donghong%20Ji%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).Neural networks for deceptive opinion spam detection: An empirical study[J].[Information Sciences](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=journaluri%3A%2829f27e6797d574e6%29%20%E3%80%8AInformation%20Sciences%E3%80%8B&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited), 2017.

1. [L Li](http://xueshu.baidu.com/s?wd=author%3A%28Luyang%20Li%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson" \t "_blank), [B Qin](http://xueshu.baidu.com/s?wd=author%3A%28Bing%20Qin%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [W Ren](http://xueshu.baidu.com/s?wd=author%3A%28Wenjing%20Ren%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [T Liu](http://xueshu.baidu.com/s?wd=author%3A%28Ting%20Liu%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).Document representation and feature combination for deceptive spam review detection[J]. [Neurocomputing](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=journaluri%3A%28738b17840f6f93f4%29%20%E3%80%8ANeurocomputing%E3%80%8B&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited), 2017.

1. [Y Yuan](http://xueshu.baidu.com/s?wd=author%3A%28Yuan%20Yuan%29%20Department%20of%20Computer%20Science%20and%20Technology%20%20Tsinghua%20University&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson" \t "_blank), [S Xie](http://xueshu.baidu.com/s?wd=author%3A%28Sihong%20Xie%29%20Department%20of%20Computer%20Science%20and%20Engineering%20%20Lehigh%20University&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [CT Lu](http://xueshu.baidu.com/s?wd=author%3A%28Chun-Ta%20Lu%29%20Department%20of%20Computer%20Science%20%20University%20of%20Illinois%20at%20Chicago&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [J Tang](http://xueshu.baidu.com/s?wd=author%3A%28Jie%20Tang%29%20Department%20of%20Computer%20Science%20and%20Technology%20%20Tsinghua%20University&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [PS Yu](http://xueshu.baidu.com/s?wd=author%3A%28Philip%20S.%20Yu%29%20Department%20of%20Computer%20Science%20%20University%20of%20Illinois%20at%20Chicago&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).Interpretable and effective opinion spam detection via temporal patterns mining across websites[C]. [IEEE International Conference on Big Data](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=confuri%3A%28dee6c6d7cd5c5fd2%29%20IEEE%20International%20Conference%20on%20Big%20Data&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited), 2017.
2. Ren Yafeng,Ji Donghong,Yin Lan.[Deceptive](http://www.cnki.com.cn/Article/CJFDTOTAL-SCLH201403010.htm) Reviews Detection Based on Semi-supervised Learning Algorithm[J].Journal of sichuan university(Enginering Science Edition),2014,l(3).

1. [H Li](http://xueshu.baidu.com/s?wd=author%3A%28Huayi%20Li%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson" \t "_blank), [G Fei](http://xueshu.baidu.com/s?wd=author%3A%28Geli%20Fei%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [S Wang](http://xueshu.baidu.com/s?wd=author%3A%28Shuai%20Wang%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [B Liu](http://xueshu.baidu.com/s?wd=author%3A%28Bing%20Liu%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson), [W Shao](http://xueshu.baidu.com/s?wd=author%3A%28Weixiang%20Shao%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson).Modeling review spam using temporal patterns and co-bursting behaviors[J].arXiv preprint [arXiv:1611.06625](https://arxiv.org/abs/1611.06625),2016.
2. Huayi Li, Geli Fei, ShuaiWang, Bing Liu.Bimodal Distribution and Co-Bursting in Review Spam Detection[C]. Proceedings of the 26th International Conference on World Wide Web,2017.

**附中文参考文献:**

[17] 吴含前,朱云杰,谢珏.[基于逻辑回归的中文在线评论有效性检测模型](http://www.cnki.com.cn/Article/CJFDTOTAL-DNDX201503004.htm)[J].东南大学学报(自然科学版),2015年03期.

[41] 任亚峰,姬东鸿,尹兰.[基于半监督学习算法的虚假评论识别研究](http://www.cnki.com.cn/Article/CJFDTOTAL-SCLH201403010.htm)[J].四川大学学报(工程科学版),2014年03期.

[20] [任亚峰](http://xueshu.baidu.com/s?wd=authoruri%3A%28a8d289bad7bbfbec%29%20author%3A%28%E4%BB%BB%E4%BA%9A%E5%B3%B0%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited),[姬东鸿](http://xueshu.baidu.com/usercenter/data/author?cmd=authoruri&wd=authoruri%3A%28a009fb14d7cf2266%29%20author%3A%28%E5%A7%AC%E4%B8%9C%E9%B8%BF%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2),[张红斌](http://xueshu.baidu.com/s?wd=authoruri%3A%288a0ca3b0d6b80b17%29%20author%3A%28%E5%BC%A0%E7%BA%A2%E6%96%8C%29%20%E6%AD%A6%E6%B1%89%E5%A4%A7%E5%AD%A6%E8%AE%A1%E7%AE%97%E6%9C%BA%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited),[尹兰](http://xueshu.baidu.com/s?wd=authoruri%3A%28660fb7648f52b237%29%20author%3A%28%E5%B0%B9%E5%85%B0%29%20%E8%B4%B5%E5%B7%9E%E5%B8%88%E8%8C%83%E5%A4%A7%E5%AD%A6%E6%95%B0%E5%AD%A6%E4%B8%8E%E8%AE%A1%E7%AE%97%E6%9C%BA%E7%A7%91%E5%AD%A6%E5%AD%A6%E9%99%A2&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson&sort=sc_cited).基于PU学习算法的虚假评论识别研究[J].计算机研究与发展,2015年03期.

**作者简介：**

稿件编号：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 伍星 | 出生年 | 1978 | 性别 | 男 | | | 籍贯 | | 四川省遂宁市 | | |  | |
| 学历  （中文 /英文） | 博士Doctor | | | | | | | | | | | |
| 职称  （中文 /英文） | 副教授 | 职务 |  | | | | CCF会员号 | | | |  | |
| 研究方向  （中文 /英文） | 自然语言处理，文本挖掘 | | | | | | | | | | | | | |
| 联系电话 | 13696449805 | | | | | E-mail | | | wuxing@cqu.edu.cn | | | | | |
| 通讯地址  （中文 /英文） | 重庆市重庆大学计算机学院  College of Computer Science,Chongqing University,Chongqing | | | | | | | | | | | 邮政编码 | | 400044 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 张圣 | 出生年 | 1992 | 性别 | 男 | | | 籍贯 | | 四川省简阳市 | | | C:\Users\zs19920826\Desktop\寸照.png | |
| 学历  （中文 /英文） | 硕士研究生Master degree candidate | | | | | | | | | | | |
| 职称  （中文 /英文） | 无 | 职务 | 无 | | | | CCF会员号 | | | | 无 | |
| 研究方向  （中文 /英文） | 自然语言处理，欺诈评论检测/[natural language processing](http://www.baidu.com/link?url=D3iP8_6r7R7TTkzj0UM9roewDJekZgDlFbZEbiUZCv9mt0p9Gu3KeJooABhSTBAnRMDhb1qNJtLOw-3e71DXAdChKYhTLjSe4W9ta3vjZnj0M99bUp4LspcNLAfo1bWHo_EqeZr_UJ38V7be25ZnNK),opinion spam | | | | | | | | | | | | | |
| 联系电话 | 17783490826 | | | | | E-mail | | | zhangsheng@cqu.edu.cn | | | | | |
| 通讯地址  （中文 /英文） | 重庆市重庆大学计算机学院  College of Computer Science,Chongqing University,Chongqing | | | | | | | | | | | 邮政编码 | | 400044 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 邹东升 | 出生年 | 1978 | 性别 | 男 | | | 籍贯 | | 河南省信阳市 | | |  | |
| 学历  （中文 /英文） | 博士后Post Doctor | | | | | | | | | | | |
| 职称  （中文 /英文） | 副教授/Associate Professor | 职务 |  | | | | CCF会员号 | | | | 32467M | |
| 研究方向  （中文 /英文） | 机器学习 数据挖掘/ Machine Learning Data Mining | | | | | | | | | | | | | |
| 联系电话 | 13883922399 | | | | | E-mail | | | dszou@cqu.edu.cn | | | | | |
| 通讯地址  （中文 /英文） | 重庆市沙坪坝区沙正街174号重庆大学A区主教1836  Room 1836, Main Building, Campus A, Chongqing University, No.174 ShaZheng Street, Shapingba District, Chongqing 400030,China | | | | | | | | | | | 邮政编码 | | 400030 |

支持基金：国家自然科学基金61309013，重庆市科技计划项目基础科学与前沿技术研究专项项目Cstc2014jcyjA40042

修改建议：如果能够在引言中增加垃圾商品评论检测问题的主要特点和难点以及需要解决的关键问题，则有利于读者在一开始阅读该文时对研究问题有一个宏观的把握和认识。

修改说明：

1. 垃圾商品评论检测的主要特点：垃圾商品评论检测作为评论情感分析的预处理，主要特点是：可用于检测的特征选择面大、可用于检测的机器学习方法多种多样、评论欺诈行为易于隐藏和改进。
2. 垃圾商品评论检测的难点：垃圾商品评论检测研究难点在于缺乏权威数据集，而人工标注商品评论的难度很大。标注数据集的质量影响了传统监督学习算法和半监督学习算法在本领域的效果。
3. 垃圾商品评论检测继续解决的关键问题：1）绝大多数现有检测技术是在单个小数据集上完成的，而特征的有效性会随着数据集变更和数据集规模增大而改变，因此，模型对大数据集的兼容性是投入实际应用的关键。2）绝大多数现有检测技术是基于静态数据的，以离线方式执行各自的欺诈检测方案。欺诈行为很可能会在垃圾评论发布过程中明显暴露，而又被之后的评论事件掩埋，因此，实时检测也是急需解决的问题。