

DOI:10.11992/tis.201603027  
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20160513.0911.004.html>

# 融合情感极性和逻辑回归的虚假评论检测方法

赵军<sup>1,2</sup>, 王红<sup>1,2</sup>

(1. 山东师范大学 信息科学与工程学院, 山东 济南 250014; 2. 山东省分布式计算机软件新技术重点实验室, 山东 济南 250014)

**摘 要:**在线购物评论为消费者比较商品的质量和其他一些购买特性提供了有用信息,然而却有大量的虚假评论者受利益驱使撰写虚假或者不公正的评论来迷惑消费者。先前的研究一般都是使用文本相似度和评分模式来探测虚假评论,这些算法可以检测特定类型的攻击者,在现实场景中许多虚假评论者刻意模仿正常用户对商品进行评论,因此先前的算法对检测这类攻击效果不佳。本文通过分析评论文本的感情极性,抽取不同的特征并使用逻辑回归模型来检测虚假评论;首先,借用自然语言处理的相关技术来分析评论文本的情感极性,判断每个用户的情感偏离大众情感的程度,如果偏离越大则说明其是虚假评论者的概率就越大;然后再选取其他几个重要特征结合逻辑回归模型进行虚假检测;通过实验对比,表明了该方法取得了较好的效果。

**关键词:**电子商务;虚假评论;购物行为;情感极性;逻辑回归

**中图分类号:**TP39    **文献标志码:**A    **文章编号:**1673-4785(2016)03-0336-07

中文引用格式:赵军,王红.融合情感极性和逻辑回归的虚假评论检测方法[J]. 智能系统学报, 2016, 11(3): 336-342.  
英文引用格式:ZHAO Jun,WANG Hong.Detection of fake reviews based on emotional orientation and logistic regression[J]. CAAI transactions on intelligent systems, 2016,11(3): 336-342.

## Detection of fake reviews based on emotional orientation and logistic regression

ZHAO Jun<sup>1,2</sup>, WANG Hong<sup>1,2</sup>

(1. School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China; 2. Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology, Ji'nan 250014, China)

**Abstract:**Online shopping reviews provide valuable customer information for comparing the quality of products and several other aspects of future purchases. However, spammers are joining this community to mislead and confuse consumers by writing fake or unfair reviews. To detect the presence of spammers, reviewer styles have been scrutinized for text similarity and rating patterns. These studies have succeeded in identifying certain types of spammers. However, there are other spammers who can manipulate their behaviors such that they are indistinguishable from normal reviewers, and thus, they cannot be detected by available techniques. In this paper, we analyze the orientation of comments, extract different features, and use a logic regression model to detect false comments. First, we utilize natural language processing technology to analyze the orientation of comments and compute the departures of those comments from those of the general public. The greater is the deviation, the greater is the probability of the comment being generated by a spammer. Then, we select several other important features and combine them with the logic regression model to identify fake comments. The experimental results verify the greater accuracy of the proposed method.

**Keywords:**Electronic commerce; fake review; shopping behavior; emotional polarity; logic regression

收稿日期:2016-03-17. 网络出版日期:2016-05-13.  
基金项目:国家自然科学基金项目(61373149,61472233);山东省科技计划项目(2012GGX10118,2014GGX101026);山东省教育科学规划项目(ZK1437B010).  
通信作者:王红.E-mail:wanghong106@163.com.

互联网的迅速崛起带来了传统商业模式的解放,传统的实体购物模式在遭遇网上购物模式时受到了巨大的冲击。电子购物网站上的商品种类齐全,同时网上购物方便快捷,可以节省消费者一定的

购物时间,迎合了当代人们快速的生活节奏;但网上购物给我们带来方便的同时也存在着自身固有的一些挑战。最具挑战性的是消费者无法像在实体店一样真实地感受到商量的质地、性能等特点,只能通过购物网站展示的图片 and 文本描述对商品有一个大致的了解。由于消费者无从得知商品地真实质量,所以他们开始过多的关注商品的评论,好的评论可以提升该产品的信誉,这会诱导更多的顾客购买该商品;相反,较差的评论无疑会降低商品的信誉值,这会大大降低该商品的成交量。正是由于这一原因,商家为了获得更高的利益,开始雇佣网络水军冒充普通顾客对自己的商品进行好评,对竞争对手的产品进行差评以达到提升自己品牌的信誉、诋毁竞争对手信誉的目的。这些广泛存在的不真实评论不仅会误导消费者的购物决策,同时也危害了电子商务经济的健康发展,为了提高消费者的购物体验,并保证市场经济的合理健康发展,因此迫切需要研究一种方法来发现并抑制虚假攻击。

### 1 相关工作

近年来,国内外学者在垃圾邮件<sup>[1]</sup>和垃圾网页<sup>[2]</sup>的识别研究上做了大量工作,并取得了较好的效果。美国伊利诺斯大学的 Bing Liu<sup>[3]</sup>教授团队于 2007 年首次提出垃圾检测,之后垃圾检测就成为了一个研究热点。

Jindal 等<sup>[4]</sup>发现商品中存在着大量的虚假评论,并且这些评论在本质上与垃圾邮件和垃圾网页截然不同,他们利用产品的评论数据,考虑评论文本、评论者和产品特征这 3 个因素进行建模来区分复制观点和非复制观点,若判断为复制观点,则将该评论归为虚假评论。

WU F 等<sup>[5]</sup>根据流行度是否被打乱来识别虚假评论。这两种方法都是基于启发式的策略,过程较为复杂。Tan 等<sup>[6]</sup>利用电阻距离来判断评论之间的上下文语义相似性,提出了一种基于电阻距离的无关虚假评论自动检测方法,该方法取得了不错的实验效果。

OTT 等<sup>[7]</sup>利用众包平台创造出了一个用于识别众包攻击的“黄金”数据集,该数据集包含真实评论和人为的虚假评论两部分,在该数据集上,他们把虚假探测问题转化为经典的文本极性分类问题解决。

任亚峰等<sup>[8]</sup>提出了一种基于语言结构和情感极性的虚假评论识别方法,从自然语言处理层面分

析评论文本的正面情感和负面情感影响,最后使用遗传算法,通过复制、交叉和变异实现种群的进化,从而提高探测准确率。

Guan 等<sup>[9]</sup>为了识别在线商店的虚假评论者提出了一种社交评论图的方法,他们提出了一种全新的概念——评论图,他们捕获了所有与某个商店相关联的评论者和评论,并把这三者构造成了一个异质网络,通过交互计算三者之间的影响来检测虚假评论,实验表明他们方法的正确探测率高达 91.24%。

Gao 等<sup>[10]</sup>为了解决网上评分系统中的虚假攻击行为,提出一种基于群组排序的虚假探测方法,他们把对相同商品评分相同的用户分为一组,根据群组的规模来检测虚假评论者,实验表明,他们所提出的方法取得了很好的探测效果。

通过总结前人已有的工作可以发现,大部分研究者要么从评论文本着手,利用自然语言处理技术,分析评论文本的正负情感极性,从而达到探测虚假评论的目的;另一种常见的方法是分析商品的评分,通过分类或者聚类算法对不同的评分进行分组,挑选出虚假的商品评分;本文提出融合情感极性和逻辑回归模型来检测虚假攻击。

### 2 评论文本情感极性分析

评论文本的情感倾向分析是通过挖掘和分析评论文本中的立场、观点、情绪等主观信息,分析出评论者的正面或者负面情感趋向。

本文主要采用基于情感知识的情感极性判别方法,通过比较文本中的正负情感词个数来判断评论文本的情感极性,情感极性判别方式为

$$\text{情感} = \begin{cases} \text{正情感词数} > \text{负情感词数} & \text{正向} \\ \text{正情感词数} = \text{负情感词数} & \text{中性} \\ \text{正情感词数} < \text{负情感词数} & \text{负向} \end{cases}$$

唐波等<sup>[11]</sup>考虑了否定词对情感词极性的影响,通过他们的实验可以看出,考虑否定词这一特征对情感倾向性分析具有重大意义。在本文中,也考虑了否定词这一重要因素,并且构建专用于评论文本分析的否定词词典,判断规则为:统计一句话中的否定词个数,若个数为偶数则该句的倾向性不变;若否定词的个数为奇数,那么语句的倾向性发生逆转。

何凤英等<sup>[12]</sup>考虑了程度副词文本的倾向性影响,提出了程度副词的 4 个量级:极量、高量、中量和低量,其对应的权重依次递减,考虑程度副词这一因素也提高了情感极性的分类精度。

邸鹏等<sup>[13]</sup>曾提出了基于转折句式的文本分析方法,他们的任务主要是基于长文本的情感分析,所以考虑上下文的转折关系是很有效的。但是他们的方法直接应用于评论文本是不合理的,因为评论文本往往是一两句话的超短文本,无从考虑上下文信息,所以本文提出了基于关联词的分析方法,把分析的单位缩小到词语级别,着重分析关联词前后的情感极性,其过程如算法 1 所示。

#### 算法 1 Review Text Orientation Analysis

输入 Review Text

输出 Review Orientation

str←{ReviewText};

While( str.read() )

For all  $i \in \text{str.Length}$  DO

IF ( str.words <sub>$i$</sub>   $\notin$  Review Dictionary ) THEN

SO←0; EXIT;

ELSE {

IF ( str.words <sub>$i$</sub>   $\in$  NegDictionary ) THEN

IF ( count % 2 == 0 ) THEN

WO <sub>$i$</sub> ←WO <sub>$i$</sub> ;

ELSE WO <sub>$i$</sub> ←-WO <sub>$i$</sub> ;

IF ( str.words <sub>$i$</sub>   $\in$  AdjDictionary ) THEN

WO <sub>$i$</sub> ←WO <sub>$i$</sub>  \* Wadj

IF ( str.words <sub>$i$</sub>   $\in$  Adversative ) THEN

WO <sub>$i$</sub> ←-WO <sub>$i$</sub> ;

END ELSE

SO =  $\sum_{i=1}^N \text{WO}_i // N = \text{str.Length}$

END FOR ;

END WHILE;

### 3 逻辑回归模型

在现实生活中经常需要分析离散变量问题,本文关心的是,哪些因素对虚假检测的影响因素更明显,这类问题实质上是一个回归问题,因变量就是上述提到的这些离散变量,自变量  $X$  是与之有关的一些因素。因为因变量是离散的,不能直接使用线性回归分析方法解决,这时最好的解决方案是 Logistic 回归模型,它对因变量的分布没有要求。与线性回归分析相似,Logistic 回归分析的基本原理就是利用一组数据拟合一个 Logistic 回归模型,然后借助这个模型揭示总体中若干个自变量与一个因变量取某个值的概率之间的关系。因变量  $Y$  是一个二值变量,取值为

$$Y = \begin{cases} 0, & \text{真实评论} \\ 1, & \text{虚假评论} \end{cases}$$

自变量为  $X_1, X_2, \dots, X_m$ ,  $P$  表示在  $m$  个自变量作用下事件发生的概率。想找到  $P$  与自变量的关系,如果用多元线性回归方程。

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p \quad (1)$$

无论  $b_i$  和  $X_i$  取任何实数,  $\ln(P/Q)$  值都可以与  $0 \sim 1$  范围内的  $P$  值相对应。

$$\ln\left(\frac{P}{Q}\right) = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p \quad (2)$$

$$P = \frac{e^{b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p}}{1 + e^{b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p}} \quad (3)$$

我们将所研究的问题转换一个角度,不是直接分析  $Y$  与  $X$  的关系,而是分析  $Y$  取某个值的概率  $P$  与  $X$  的关系。令  $Y$  为 1、0 变量,  $Y=0$  表示正常评论,  $Y=1$  表示虚假评论,  $X$  是与虚假评论有关的因素。如果  $P$  表示虚假攻击的概率,那么研究虚假攻击的概率  $P$  与  $X$  有关因素的关系就相对简单一些。

#### 3.1 变量选取的衡量标准

优势比 OR(odds ratio)是流行病衡量危险因素作用大小的比数比例指标,其计算公式为

$$\text{OR}_j = \frac{P_1 / (1 - P_1)}{P_0 / (1 - P_0)} \quad (4)$$

式中:  $P_1$  和  $P_0$  分别表示在  $X_j$  取值为  $c_1$  及  $c_0$  存在虚假攻击的概率,  $\text{OR}_j$  称作多变量调整后的优势比,表示扣除了其他自变量影响后的危险因素的作用。对比某一个危险因素两个不同暴露水平  $X_j = c_1$  与  $X_j = c_0$  的虚假攻击情况(假设其他因素的水平相同),其优势比的自然对数为式(5):

$$\ln \text{OR}_j = \ln \left[ \frac{P_1 / (1 - P_1)}{P_0 / (1 - P_0)} \right] = \text{logit} P_1 - \text{logit} P_0 \quad (5)$$

$$\text{若 } X_j = \begin{cases} 1, & \text{虚假评论} \\ 0, & \text{正常评论} \end{cases}$$

$$c_1 - c_0 = 1$$

则

$$\begin{aligned} \text{OR}_j &= \exp \beta_j \\ \beta_j &\begin{cases} = 0, & \text{OR}_j = 1 \quad \text{无作用} \\ > 0, & \text{OR}_j > 1 \quad \text{虚假评论} \\ < 0, & \text{OR}_j < 1 \quad \text{正常评论} \end{cases} \quad (6) \end{aligned}$$

#### 3.2 变量选择

使用逻辑回归模型时,主要有 3 种选择变量的方式:前向选择、后向选择和逐步回归。本文采用逐步回归的方式进行变量选择,其基本思想是逐个引

入影响模型的自变量,每次都是引入对  $Y$  影响最为显著的自变量,并对方程中存在的变量逐个进行检验,把变为不显著的变量逐个从模型中删除,最终,使得模型中存在的变量是对  $Y$  影响最为显著的变量,筛选的步骤如下:首先给出引入变量的显著性水平  $\alpha_{in}$  和剔除变量的显著性水平  $\alpha_{out}$ ,然后按下图 1 进行筛选,筛选过程如下。

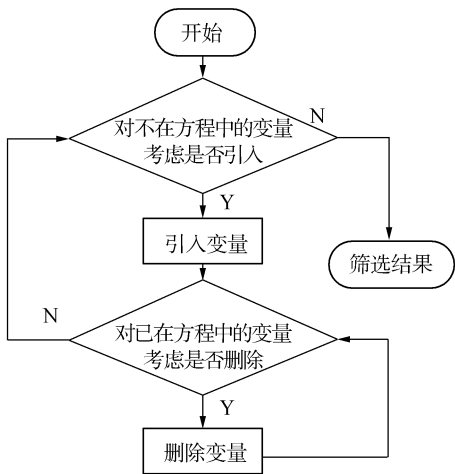


图 1 变量筛选流程图

Fig.1 Flow chart of variable selection

1) 设仅有截距项的最大似然估计值为  $L_0$ , 分别计算  $m$  个变量的 Score 检测值, 设有最小  $P$  值的变量为  $X_{e1}$ , 且有  $P_{e1} = \min(P_j)$ , 若  $P_{e1} < \alpha_{in}$ , 则此变量进入逻辑回归模型, 不然停止。如果此变量是名义变量单面化的变量, 则此变量的其他单面化变量也会进入模型, 其中  $\alpha_{in}$  表示要引入变量的显著性水平。

2) 为了确定当变量  $X_{e1}$  在模型中时其他  $m-1$  个变量是否也重要, 将  $X_j, j = 1, 2, \dots, m, j \neq e$  分别与  $X_{e1}$  进行拟合。对  $m-1$  个变量分别计算 Score 检测值, 其  $P$  值设为  $P_j$ 。设有最小  $P$  值的变量为  $X_{e2}$ , 且有  $P_{e2} = \min(P_j)$ 。若  $P_{e2} < \alpha_{in}$ , 则进入下一步, 否则停止迭代。

3) 此时的模型中已经包含  $X_{e1}$  和  $X_{e2}$  两个变量, 但是我们应该注意到, 在引入变量  $X_{e2}$  之后, 变量  $X_{e1}$  可能不再是显著性变量, 我们应该使用 Wald 检验分别计算它们的显著值和响应的  $P$  值。如果  $P$  值大于  $\alpha_{out}$ , 则此变量从模型中删除, 否则停止迭代计算;

4) 依次进行迭代计算, 每当向前选择一个变量进入后, 都进行向后删除的检查, 循环终止的条件是: 模型中的所有自变量的  $P$  值都小于  $\alpha_{out}$ , 被删除自变量的  $P$  值都大于  $\alpha_{in}$ 。

4 评论文本特征工程

这里首先提出可能影响虚假评论的 10 个特征, 然后使用逐步回归方法计算哪个特征对其影响严重, 10 个特征如下。

**文本长度 ( $F_1$ )** 统计评论文本中的文字个数, 正常用户一般懒于评论只给出评分或给出简短的评论文本, 而虚假评论者为了提高或贬低某件商品, 需要写下比较冗长的评论文本。

**复杂度 ( $F_2$ )** 评论文本的复杂度是指文本中所使用词汇的复杂度, 一般认为正常用户所使用的评论词汇比较简单, 而虚假评论的词汇相对比较复杂和专业。

**关联度 ( $F_3$ )** 评论关联度是指评论文本与被评论商品的相关程度, 有些虚假评论者为了完成评论字数要求, 常常会复制一些与商品无关的文本来完成虚假评论任务。

**一致性 ( $F_4$ )** 一致性评论文本的情感强度与所给评分的相似度, 是探测随机评论的重要特征。

**情感强度 ( $F_5$ )** 情感强度是指评论文本中的情感极性强度, 普通用户的评论情感比较公正, 而虚假评论的情感强度比较激烈。

**是否包含转折词 ( $F_6$ )** 正常用户可能对产品的部分性能指标是满意的, 而对另外的某些特征是不太满意的, 所以他们的评论中常常包含转折词, 而虚假评论者的情感极性高度一致, 很少会有情感的逆转。

**复制文本 ( $F_7$ )** 虚假评论者为了快速完成自己的虚假攻击任务, 往往会把某个商品的虚假评论复制对商品以不同的消费者身份进行评论。

**用户信誉 ( $F_8$ )** 某个用户的信誉是由他所发表的评论被其他消费者采纳的数量决定的, 如果采纳该用户的人数越多, 说明其信誉越高。

**初评和追评的一致性 ( $F_9$ )** 正常评论者在初评时一般只是从商品的外观给商品进行评分, 对该商品使用一定时间后, 可能会给出更加具有意义的追评; 而虚假评论者的初评和追评几乎是同时完成的。

**附广告图片 ( $F_{10}$ )** 现在商家为了提高商品信誉, 还会进行“好评晒图返现”活动, 普通正常用户很少拍图上传, 而虚假评论中往往会伴随着图片的出现, 我们认为评论附图的评论很可疑。

5 特征选择

本文主要使用优势比 (OR) 和逐步回归变量筛



选方法来选择对逻辑回归模型影响最为显著的特征。

5.1 数据集

在 Liu<sup>[4]</sup>从 Amazon 收集并整理的 415 179 条评论数据的基础上,筛选出含有评论文本的 13 246 条评论作为本文的实验数据,其中包含虚假评论 3 412 条,真实评论 9 834 条,实验数据集如表 1。

表 1 实验数据表  
Table 1 Experimental data table

| 数据类型 | 条数     |
|------|--------|
| 虚假评论 | 3 412  |
| 正常评论 | 9 834  |
| 评论总数 | 13 246 |

5.2 计算候选自变量的优势比

为了检测哪些候选自变量对逻辑回归模型的影响更为显著,本文使用式(4)计算每个自变量的优势比,以评论文本长度为例来说明优势比的计算步骤。

表 2 文本长度分布表  
Table 2 The distribution of text length

| 结果   | 文本长度≥50 | 文本长度<50 | 合计     |
|------|---------|---------|--------|
| 虚假评论 | 2 536   | 876     | 3 412  |
| 正常评论 | 1 458   | 8 376   | 9 834  |
| 合计   | 3 994   | 9 252   | 13 246 |

文本长度≥50 的优势= $\frac{2\,536/3\,994}{1\,458/9\,994}\approx 1.74$

文本长度 <50 的优势= $\frac{876/9\,252}{8\,376/9\,252}\approx 0.10$

则  $OR = \frac{1.74}{0.10} = 17.4$ , 所以该特征对检测虚假评

论的影响是非常显著的。特征的优势比如表 3。

表 3 10 个候选特征及 OR 值

| 变量       | 赋值说明       | OR   |
|----------|------------|------|
| $F_1$    | 文本长度≥50=1  | 17.4 |
|          | 文本长度<50=0  |      |
| $F_2$    | 复杂=1,不复杂=0 | 5.31 |
| $F_3$    | 不关联=1,关联=0 | 0.34 |
| $F_4$    | 不一致=1,一致=0 | 3.87 |
| $F_5$    | 强烈=1,不强烈=0 | 7.68 |
| $F_6$    | 不含=1,包含=0  | 15.6 |
| $F_7$    | 是=1,不是=0   | 2.87 |
| $F_8$    | 高=1,低=0    | 0.78 |
| $F_9$    | 不一致=1,一致=0 | 7.46 |
| $F_{10}$ | 附图=1,不附图=0 | 4.23 |

5.3 模型检验

为了判断选取的模型是否有效,需要对该模型进行检验。逻辑回归模型主要包含 3 种假设检验方法,分别是似然比检验、Wald 检验和计分检验,其中,似然比检验既适用于单个影响因素的假设检验,又适用于多个影响因素的同时检验;Wald 检验适合单个影响因素的检验;计分检验与传统的 Mantel-haenszel 检验结果相同,在小样本空间中比似然比检验更接近 $\chi^2$ 分布;本文采用似然比检验方法对模型进行检验。

似然比检验的原理是通过分析模型中变量变化对似然比的影响,依此来判断增加或者去除某个自变量是否对因变量有显著影响,检验统计量的公式如式(7)所示。

$G = -2[\ln(L_{m-1}) - \ln(L_m)]$  (7)

式中: $\ln(L_{m-1})$ 为不包含检验变量时模型的对数似然值, $\ln(L_m)$ 为包含检验变量时模型的对数似然值.当检验一个变量时  $G$  服从自由度为 1 的 $\chi^2$ 分布,当对整个模型进行检验时,就是相当于模型中所有的  $m$  个变量的回归系数为 0, $G$  服从自由度为  $m$  的 $\chi^2$ 分布。如果检验结果为拒绝  $H_0$ ,则表示该影响因素对回归模型有统计学意义,即对事件发生有影响。我们分别计算了所选的 6 个特征的似然比,在显著水平  $\alpha = 0.05$  的条件下,计算的结果如表 4 所示。

表 4 似然比测试表  
Table 4 Likelihood ratio test

| 变量  | $F_1$ | $F_2$ | $F_5$ | $F_6$ | $F_9$ | $F_{10}$ |
|-----|-------|-------|-------|-------|-------|----------|
| $G$ | 21.5  | 7.3   | 11.5  | 14.5  | 13.6  | 8.2      |

因为这 6 个变量的检验统计量都大于 $\chi^2_{0.05,1} = 3.84$ ,所以可以得出如下结论,在其他 5 个变量不变的情况下,第 6 个变量对模型的影响是比较显著的,所以选取的这 6 个变量对虚假攻击检测具有一定的指导意义。

6 实验分析

本文借用 Spss 统计工具对实验数据进行分析,采用对比实验的方式验证本文方法的有效性,由于数据样本有限,我们采用五折交叉验证,分别与任亚峰等<sup>[8]</sup>提出的基于语言结构和情感极性的虚假评论识别方法和 Feng 等<sup>[14]</sup>提出的基于句法结构的方

法进行对比,从而分析本文方法的优势和不足。本文采用最为通用的 3 个评判指标来判断虚假检测的优劣,即准确率、召回率和  $F_1$  值。从图中可以发现融合情感倾向和逻辑回归模型的虚假攻击检测方法表现出了不错的性能。

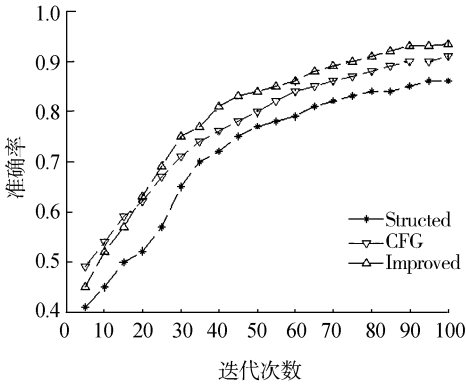


图 2 准确率对比  
Fig.2 Accuracy comparison

任亚峰等认为传统的启发式策略或者全监督学习方法很难有效地解决虚假攻击的检测问题,他们分析了虚假评论和真实评论之间在语言结构和情感极性的差异,借用经典的遗传算法对文本的书写结构和情感极性两大主要特征进行优化选择,最终选择出了 5 个最为重要的特征,最后又根据这些特征结合无监督的硬聚类和软聚类算法实现了虚假评论的检测。该方法的检测准确率有所提高,主要原因是在考虑情感极性和强度的同时,融合了评论文本的较为重要的 5 个特征,使其分析的更加全面。该方法也有不足之处,即没有突出哪些特征最为重要,如果能再把不同特征的权重考虑进去,检测的效果可能会更好。召回率对比和  $F_1$  指数对比见图 3 和图 4。

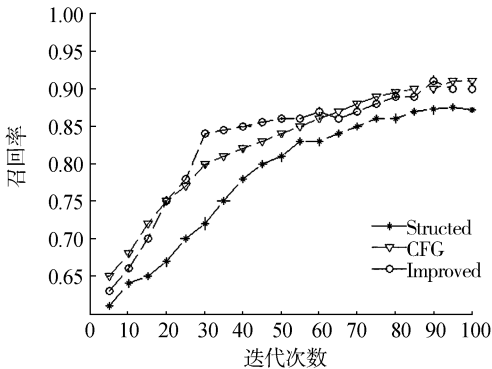


图 3 召回率对比  
Fig.3 Recall comparison

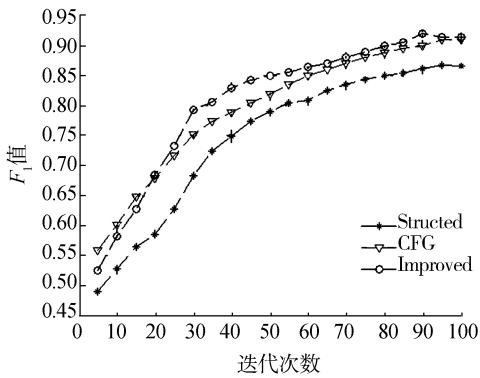


图 4  $F_1$  指数对比  
Fig.4  $F_1$ -Measure comparison

Feng 等提出的基于句法结构的虚假评论检测算法相比于任亚峰等提出的算法又有明显的提高,他们认为浅层次的句法模式是不可靠的,他们的工作主要研究了深层次的句法模式,并在前人的研究基础之上加入了一些非常规的句法模式来构建语义树。他们又从 4 个不同数据集中利用文本句法树提取了几种不常见的语义特征,该方法使其检测精度达到了 91.2%,召回率也有明显提高。我们分析工作性能提高的主要原因是,购物评论往往是不规则的,文本长度极短并且没有固定的书写模式,所以任亚峰等提出的基于简单语法模式的检测算法受到了一定的限制,而 Feng 等正是针对评论文本和常规文本在句法上存在明显差异的前提下,利用语义树,挖掘深层的句法关系,构建出了专门用于分析评论短文本的语义树,使得检测性能得到大幅度提高。

本文方法相对于二者的研究方法准确率有所增加,而召回率比 Feng 等提出的方法的要差一些。本文方法首先提出了基于关联词(特别是转折词)的情感极性判别方法,以此来判断评论的情感是否出现转变、情感强度是否异常极端,然后又提取了 10 个候选的评论文本特征,经过极大似然检验后,选取了 6 个最为重要的特征,最后使用逻辑回归模型对评论进行检测,促使检测的准确率有所增大。但是在加入更多特征使检测精度提高的同时,我们也发现,召回率比 Feng 等的要低一些,原因可能是更多的特征被加入了模型中,对数据的质量提出了更高的要求,如一些文本中可能不存在关联词这一特征。

7 结束语

随着电子商务的蓬勃发展,研究者们对虚假评论检测作出了不懈的努力。针对评论文本过短,评论随意等特点本文提出了融合关联词的情感倾向分析,然后再此基础上又分析了影响分类的 6 个特征,把这些特征应用于逻辑回归模型中实现了虚假评论

的检测,该方法取得了不错的效果。但该模型还有待改进,首先,文本的极性判别方式过于粗糙,只是考虑了文本中的否定词和关联词,没有特别注重分析程度副词,只分析了文本的极性,却没有定量分析其情感强度;其次,模型中没有动态地考虑评论情况,研究表明,不诚实的商家往往在开办网店的初期雇佣虚假评论人员通过刷单的方式提高自己的信誉,所以虚假评论往往发生在电商经营的初期。最后,可能评论文本中还隐藏着许多其他的因素可以提高检测的精度,下一步的工作将主要集中在这 3 个方面。

## 参考文献:

- [1] KOLCZ A, ALSPECTOR J. SVM-based filtering of E-mail spam with content specific misclassification costs[C]//Proceedings of ICDM-2001 Workshop on Text Mining. Dallas, USA, 2001: 324-332.
- [2] BECCHETTI L, CASTILLO C, DONATO D, et al. Link-based characterization and detection of web spam[C]//Adversarial Information Retrieval on the Web. Washington, USA, 2006: 1012-1021.
- [3] JINDAL N, LIU Bing. Review spam detection[C]//Proceedings of the 16th International Conference on World Wide Web. Alberta, Canada, 2007: 1189-1190.
- [4] JINDAL N, LIU Bing, et al. Opinion spam and analysis [C]//Proceedings of the 2008 International Conference on Web Search and Data Mining. California, USA, 2008: 219-230.
- [5] WU Fang, HUBERMAN B A. Opinion information under costly express[J]. ACM transactions on intelligence systems and technology, 2010, 1(1): 5.
- [6] 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, 168.  
TAN Wentang, ZHU Hong, GE Bin, et al. Method of review spam detection[J]. Journal of national university of defense technology, 2012, 34(5): 153-157, 168.
- [7] OTT M, CHOI Y, CARIDIE 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. Portland, USA, 2011, 1: 309-319.
- [8] 任亚峰, 尹兰, 姬东鸿. 基于语言结构和情感极性的虚假评论识别[J]. 计算机科学与探索, 2014, 8(3): 313-320.  
REN Yafeng, YIN Lan, JI Donghong. Deceptive reviews detection based on language structure and sentiment polarity [J]. Journal of frontiers of computer science and technology, 2014, 8(3): 313-320.
- [9] WANG Guan, XIE Sihong, LIU Bing, et al. Identify online store review spammers via social review graph[J]. ACM Transactions on intelligent systems and technology, 2012, 3(4): 61.
- [10] GAO Jian, DONG Yuwei, SHANG Mingsheng, et al. Group-based ranking method for online rating systems with spamming attacks[J]. EPL (europhysics letters), 2015, 110(2): 28003.
- [11] 唐波, 陈光, 王星雅, 等. 微博新词发现及情感倾向性判断分析[J]. 山东大学学报:理学版, 2015, 50(1): 20-25.  
TANG Bo, CHEN Guang, WANG Xingya, et al. Analysis on new word detection and sentiment orientation in Microblog[J]. Journal of Shandong university: nature science, 2015, 50(1): 20-25.
- [12] 何凤英. 基于语义理解的中文博文倾向性分析[J]. 计算机应用, 2011, 31(8): 2130-2133, 2137.  
HE Fengying. Orientation analysis for Chinese blog text based on semantic comprehension[J]. Journal of computer application, 2011, 31(8): 2130-2133, 2137.
- [13] 邸鹏, 李爱萍, 段利国. 基于转折句式的文本情感倾向性分析[J]. 计算机工程与设计, 2014, 35(12): 4289-4295.  
DI Peng, LI Aiping, DUAN Ligu. Text sentiment polarity analysis based on transition sentence[J]. Computer engineering and design, 2014, 35(12): 4289-4295.
- [14] FENG Song, BANERJEE R, CHOI Y. Syntactic stylometry for deception detection[C]//Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2. Jeju, Korea, 2012: 171-175.
- [15] LI Jiwei, CARDIE C, LI Sujian. TopicSpam: a topic-model-based approach for spam detection[C]//Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. Sofi, Bulgaria, 2013: 217-221.
- [16] JINDAL N, LIU Bing, LIM E P. Finding unusual review patterns using unexpected rules[C]//Proceedings of the 19th ACM International Conference on Information and Knowledge Management. Ontario, Canada, 2010: 1549-1552.
- [17] JO Y, OH A H. Aspect and sentiment unification model for online review analysis[C]//Proceedings of the 4th ACM International Conference on Web Search and Data Mining. New York, USA, 2011: 815-824.

## 作者简介:



赵军,男,1989年生,硕士研究生,主要研究方向为大数据、数据挖掘、机器学习。



王红,女,1966年生,教授,博士生导师,主要研究方向为大数据、复杂网络、数据挖掘。主持国家自然科学基金项目 1 项,参与国家自然科学基金项目 3 项,主持省级基金项目 6 项,发表学术论文 43 篇。