

酒店顾客在线评论的文本挖掘

熊 伟 郭扬杰

(华南师范大学旅游管理系 广东广州 510631)

摘 要: 酒店在线评论是顾客对酒店服务质量的真实感知,学术界已有利用其进行酒店服务质量研究的相关成果,但多采用内容分析法,不能对评论进行批量处理,成果应用大受限制。基于此,本文利用文本挖掘技术的批量处理功能,选取高档商务型酒店的顾客在线评论作为研究对象,对酒店顾客的在线评论进行词频和情感分析,并结合时间序列分析法预测未来的发展情况。主要结论有:第一,近年来高档商务酒店的商务顾客对其服务质量在线评论最主要的关注面是“总体服务”和“客房”,并从技术质量向功能质量转移;第二,近年来高档商务酒店顾客对其服务质量在线评论的情绪心理以积极情绪为主且有不断加强的趋势;第三,时间序列分析结果显示,2012年上半年商务顾客对广州市高档商务酒店服务质量的关注程度从高到低依次为“总体服务”、“方便性”、“酒店环境”、“交通”和“客房”,同时,2012年上半年广州市高档商务型酒店的顾客评论中会有较多带有消极情感的评论出现。最后,笔者也对本文存在的局限性和后续可以深入的研究进行了讨论。

关键词: 文本挖掘; 服务质量; 在线评论; 商务型酒店

[中图分类号] F59

[文献标识码] A

[文章编号] 1003-6539(2013)11-0038-10

A Study of Text Mining on Online Reviews of Luxury Business Hotels in Guangzhou

Xiong Wei / Guo Yangjie

(Department of Tourism Management, South China Normal University, Guangzhou 510631, China)

Abstract: Online Hotel Reviews are an authentic reflection of hotel guests' perception of hotel service. Despite the existence of some research results on this issue in academia, the majority of them made use of content analysis, which leads to a restricted application of these findings in practice due to its limited ability of batch processing. This study aims to analyze online customer reviews of luxury business hotels in Guangzhou in terms of word frequency and emotion with the combination of text mining and time series analysis methods, in order to predict its development tendency. The research outcomes show that “overall services” and “room” are gaining increasingly attention within online reviews of luxury business hotels by business guests recently. Reviews on service quality are transferring from technical quality to functional quality as well. Secondly, positive emotion dominates online business guest reviews on service quality of luxury business hotels over the years. And its dominance becomes strengthened. Thirdly, the results of time series analysis reveal that in the first half of 2012, “overall services” and “convenience” would be concerned the most within online business guest reviews of luxury business hotels in Guangzhou. In the meantime, more reviews would be accompanied with negative emotions. Limitations and future research suggestions are also discussed.

Key words: text mining; service quality; online reviews; luxury business hotel

一、引言

服务质量管理是酒店管理的重要内容,好的

服务质量管理,可提升酒店自身的竞争力。近年来,中国酒店业迅猛发展,而广州作为商务型旅游城市,商务型酒店数量众多,尤其是在一年两届的中国进出口商品交易会(Import and Export

[基金项目] 本研究受广东省自然科学基金项目(9451063101002226)、华南师范大学青年教师科研培育基金的资助。

[收稿日期] 2013-04-13

[作者简介] 熊伟(1981~),女,湖南长沙人,硕士,华南师范大学旅游管理系教师,主要研究方向为酒店管理与旅游电子商务等。

郭扬杰(1990~),男,广东汕头人,华南师范大学旅游管理系2008级本科生,主要研究方向为旅游管理。

Fair, 简称广交会)期间,广州市的商务型酒店承担着接待来自世界各地商务客人的职能,尤其受到来自全球的关注,其服务质量的好坏直接反映了我国旅游接待服务水平的高低。酒店在线评论是顾客入住酒店后对其所感受到的服务质量的真实反映,对其进行分析可发现酒店服务质量管理存在的问题,有利于酒店有针对性地采取相关措施。

本文通过文本挖掘技术对大量网友客观真实的评价进行分析,以发现顾客对高档商务型酒店评论的关注面和情感状态的一般规律,并利用时间序列分析法预测未来的发展情况。从理论的角度而言,本文是采用新技术对酒店在线评论进行分析的新尝试,从实践角度而言,可更科学地利用酒店顾客的在线评论,为商务型酒店未来的服务质量管理提供一定的决策参考。

二、国内外研究现状

1. 服务质量评价研究

服务质量的研究始于20世纪80年代。作为企业管理重要内容的服务质量,用来描述所定制的服务提供的程度^[1]。它包括技术质量和功能质量,技术质量又称为结果质量,是指服务过程的产出质量,即顾客从服务过程中所获取的实际产出,即企业为顾客提供的服务结果的质量;功能质量是指顾客如何得到服务的,具体表现在服务接触的过程中。服务质量取决于消费者接受服务前对服务的期望与接受服务时实际感知到的服务水平之间的差距,即:服务质量=服务感知-服务期望^[2],这一定义得到了众多学者的肯定。而服务质量评价方法主要分为基于事件(Incident-based)和基于属性(Attribute-based)的两种评价方法^[3]。

关键事件技术(Critical Incidents Technique,简称CTI)是服务质量评价中基于事件评价的代表方法,通过这一方法可找出在服务接触过程中促使顾客满意或不满意的特定事件和行为,进而为形成满意度监测计划、设计服务程序与策略以及培训一线员工提供可靠的依据^[4]。

基于属性,服务质量可通过可靠性(reliability)、反应性(responsiveness)、保

证性(assurance)、移情性(empathy)和有形性(tangibles)5个基本要素来评价,在此基础上建立的SERVQUAL模型得到了广泛的应用^[5]。但此模型在概念化和操作化方面有一定的缺陷,因此有学者在此基础上提出了SERVPERF模型^[6]。

国内在服务质量评价方面的研究,一直处于引进学习阶段,在服务质量理论方面的独创性研究成果相对较少。汪纯孝在医院服务质量的实证研究中,提出服务质量包括环境质量、技术质量、感情质量、关系质量和沟通质量的结论^[7];在服务行业中,享受到服务的消费者与提高服务的企业会发生多层次、多维度的交互作用,交互过程的好坏直接影响顾客对服务的评价,决定服务质量的高低^[8]。其他大部分相关研究主要是基于现有理论所进行的实证研究^[9-13]。

2. 酒店服务质量评价研究

20世纪90年代初,国内外学者开始利用这些方法进行酒店行业服务质量评价的研究。

实验结果证实,CIT方法可运用于酒店服务质量评价^[14],其优点在其着眼于服务过程中的情节,有利于发现出现问题的具体环节。吉伯特(Gilbert)和莫里森(Morris)利用其研究了酒店与旅行安排对商务旅游者满意度的影响,认为研究关键事件能够更深刻地理解导致商务旅游者满意和不满意的因素^[15]。

在国内研究中,王文君^[16]等和苏志平^[17]分别利用SERVQUAL量表对影响酒店服务质量的因素进行了实证研究。而温碧燕所采用的约翰斯顿(Johnston)的软硬质量模型在计量顾客感知的服务质量时,不仅涵盖了SERVQUAL量表的所有内容,还包括了没有涉及的因素,如敬业、友好、美观等因素^[18]。

从整体上看,当前的酒店服务质量评价研究主要依靠服务质量反馈量表进行采样,而随着顾客的网络信息需求量的加大,酒店服务质量的研究采样开始往在线网络转移^[19],如酒店顾客在线评论,这些评论是顾客有感而发的,客观且真实^[20],可作为酒店服务质量评价研究新的采样方式。

3. 酒店顾客在线评论研究

目前,酒店在线评论的研究主要集中在评论信息挖掘和评论影响研究两个方面。

在国外,监督分类方法(supervised classification)被用于对tripadvisor^①网站内的酒店在线评论进行分类,为搜索者提供最受推荐的相关评论,并由此发现顾客对酒店服务质量的关注焦点^[21]。同时,在线评论会影响顾客决策,好评与差评对顾客选择酒店的决策影响程度更大于酒店的声望^[22]。

在国内,顾客在线评论的内容分析结果显示,三至五星级酒店的顾客对总台接待、客房服务、房间大小与酒店星级紧密相关的因素均非常重视^[20],经济型酒店的顾客对客房、噪音和网络等因素的关注度较高,且东部城市的经济型酒店综合服务质量优于中西部^[23]。这些评论的内容质量、时效性、数量、细节和评论者的威望值对商务顾客的预订意向有显著正向影响^[24]。

由此可见,当前的酒店顾客在线评论得到了国内外众多学者的关注,但依然存在两个比较突出的问题,一是在各类型各地区等细化维度上的分析或对比研究较少,而考虑到酒店须为不同顾客提供个性化服务,这种分析具有重要的实践价值;二是多数研究的方法为内容分析法,不能对评论进行批量处理,成果应用大受限制。

因此,本研究将文本挖掘技术引入到酒店顾客在线评论研究中,尽量避免内容分析法的主观性并实现批量化处理。研究选定广州市高档商务型酒店顾客在线评论为研究对象,进行词频分析与情感分析,了解顾客的关注内容和情绪心理,再通过时间序列分析法预测顾客评论在未来的关注内容与情绪变化,具有一定的理论意义和实践价值。

三、研究方法的过程

1. 研究方法

本文的研究方法主要是文本挖掘,其属于数据挖掘的范畴,是在非结构化的文本数据中抽取有价值的信息,其中包括对大量文档集合的内容

进行预处理、特征提取、文本检索、文本分类、文本聚类、文本情感分析等过程^[25],是一种比较新兴且趋于成熟的分析技术。

文本挖掘技术比较成熟的应用领域是新闻业和医学,在旅游研究领域的相关成果较少,但近年来,情感分析在旅游领域的研究受到了学者们的特别关注^[26]。如在国外,采用情感分类方法分析旅游博客中的信息,以分析旅游者对某个目的地的整体评价^[27];在国内,基于LingPipe方法和PMI方法分析携程网旅游评论的情感倾向,从中得到有价值的且有利于顾客决策的评论^[28]。也有学者对评论进一步细化分析,通过抓取携程网社区中的评论数据进行情感分析,从旅游6要素的维度出发探讨游客的体验感受。总之,目前情感分析技术在旅游行业中的应用并不深入,局限在文本挖掘的信息访问的信息浏览层面,未能应用于数据预测方面,在酒店业的相关研究中,这种应用更是有限^[26]。

2. 研究过程

(1) 样本选取

在国内,为酒店提供在线评论且具有较高知名度的网站主要有到到网^②、携程网^③、去哪儿网^④等。笔者通过对比最终选择携程网作为本次研究的样本网站。原因有二:

第一,Alexa排名^⑤是目前常引用的用来评价某一网站访问量的一个指标,也反映了网站的知名度、用户活动程度等情况^⑥。三大网站在ALEXA网站的全球排名依次为第6471名、第759名、第1126名(更新于2012年2月27日),可见携程网的知名度和用户活跃度相对较高。

第二,本次研究的对象是商务型酒店的商务顾客评论(以下简称顾客评论)。笔者对各网站进行细致浏览后发现,对于顾客出行动机的分类,携程网有包括商务型在内的5种分类体系,而其他网站中的分类不够明显。考虑到选样的针对性,携程网作为样本网站最为合适。

随后,本研究共有样本酒店17家,主要基于地理分布和品牌两大因素进行选样。一方面,广州星级酒店的发展主要集中于荔湾区、越秀区、天河区^[29],这四大区域不仅商务型酒店分布密集,而且还包括火车站、机场、CBD、

会展中心等一般商务出行涉及的场所。因此,在地理分布上,样本酒店的选取主要集中在这四大区域内。另一方面,样本涵盖了国际酒店集团和国内酒店集团旗下不同品牌的四星级和五星级酒店。

接下来,笔者选取了携程网上的样本酒店在2009年1月1日至2011年12月31日共36个月份的2130条顾客评论,在随机取样的基础上保证每个月的样本评论数量均在50条左右,并对无效评论(评论中大部分为符号或混乱的词语、语义不清晰)或重复评论(评论中出现多次语句重复)进行了手动剔除。

(2) 分析工具选取

本研究首先采用ROSTCM 6.0软件对顾客评论进行文本挖掘。ROSTCM 6.0是一款文本内容挖掘软件,集成了中文词频分析、情感分析等多种挖掘功能,中文词频分析功能可以根据使用者自定义的词典统计一个文档中的所有词及词频情况即顾客评论关注面;情感分析功能可以对文档内容的积极、中性、消极词语进行赋值计算,反映文档内容的情感情况。此外,还通过SPSS的时间序列分析,预测未来酒店在线评论中顾客的关注面与情感情况,最终实现在线评论的文本挖掘,在非结构化的文本数据中抽取有价值的信息。

在进行分析前,需要对获取的评论样本进行文本预处理,一方面防止分析中出现混乱等错误,另一方面,由于酒店评论者在使用文字

上存在不同的习惯,部分评论会出现繁体字,预处理可以保证词频分析软件能准确识别文本内容。在具体的操作中,笔者使用了Office Word 2010软件把繁体中文内容转换为简体中文,把修正后的文档全部另存为词频分析软件能够识别的txt文件。

四、广州市高档商务型酒店顾客评论的文本挖掘

1. 顾客评论关注面分析与预测

(1) 顾客评论关注面分析

在词频分析中,为了保证分词的准确性,分词词表的建立非常重要。本研究在正式进行数据分析之前,随机选取了200条评论作为试用样本,将顾客提及的关于酒店领域的词汇纳入词表,同时将酒店专业词汇(选自于《饭店管理》^①各章节主要专业词汇)补充进分词词表,建立可用于酒店评论分词的词表。最后,运用软件的过滤功能,过滤无关词汇,获取顾客评论中高频特征词及其排序情况。

通过对高频特征词表的归纳修正(由于原始高频特征词表中会出现一些顾客评论习惯添加的酒店领域词语,如酒店、入住等,笔者对此进行了进一步修正归纳),最后得到以下顾客评论的关注面与排名情况(鉴于篇幅所限,只选取排名前10的关注面),具体如表1所示。

表1 顾客评论的关注面与排名情况

排名	2009~2011年	2011年	2010年	2009年
1	总体服务	总体服务	客房	总体服务
2	客房	客房	总体服务	客房
3	交通	方便性	方便性	酒店环境
4	方便性	交通	酒店环境	硬件设施
5	酒店环境	硬件设施	硬件设施	方便性
6	硬件设施	酒店环境	交通	交通
7	价格	价格	价格	价格
8	早餐	早餐	早餐	早餐
9	酒店位置	酒店位置	前台	推荐情况
10	前台	前台	推荐情况	酒店位置

从整体上看，顾客最为关注的是总体服务，然后依次是客房、交通、方便性、酒店环境、硬件设施、价格、早餐、酒店位置和前台。客房和总体服务是顾客评论中两个最重要的关注焦点，而环境、硬件、交通等便利性问题也是顾客评论中的重要主题。相对而言，价格、早餐和前台等问题在顾客评论中的敏感程度较低。

有研究表明，客房硬件、酒店服务和酒店餐饮是商务型酒店服务质量出现问题的主要方面，而客房的评价比例达到了20%^[30]。同时，2010年《中国商务人士旅行调查研究报告》^⑧中也提及商务旅客在选择酒店中会更多考虑其交通便捷的因素。本研究的发现对此进行了验证。

总体上看，顾客评价的关注面首先具有全局性特点，大多从总体上关注酒店服务，并且集中于酒店的核心产品——客房；其次，顾客评论的关注面多为交通服务、方便性等，这些属于服务质量中的功能质量，具体表现在服务接触的过程中。因此，顾客在线评论中对其功能质量的关注

呈逐年加强的趋势。

(2) 顾客评论关注面预测

为了实现文本挖掘数据预测的应用，笔者将对顾客评论关注面的情况进行时间序列分析，预测2012年上半年样本酒店顾客评论关注面的变化情况。

时间序列分析是通过展示研究对象在一定时期内的变动过程，从中寻找和分析事物的变化特征、发展趋势和规律^[31]。本研究以月份为周期进行时间序列分析，预测2012年1月到6月顾客评论关注排名前10的变量情况（为方便说明，以下统一称之为变量）。

利用SPSS统计分析软件中的专家建模器自动为定义成功的时间序列进行预测。模型时间序列分析是建立在序列平稳的条件上的，一个平稳的随机过程应符合以下要求：均数不随时间变化；方差不随时间变化^[31]。从图1可知，“客房”变量在2009年1月1日至2011年12月31日36个月内的整个排名曲线并没有明显的变化趋势，属于平稳时间序列，达到预测模型平稳性的要求，可直接进行时间序列分析。

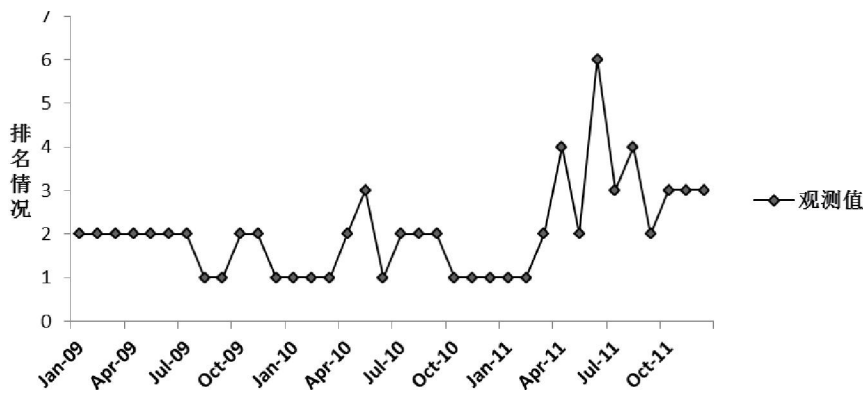


图1 “客房”变量的排名曲线情况

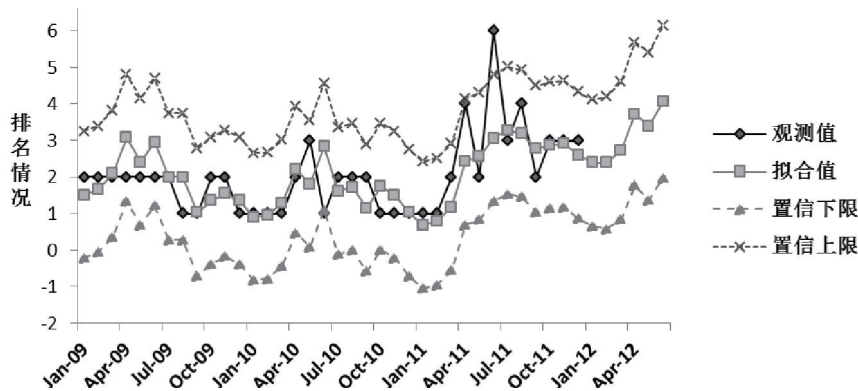


图2 “客房”变量的预测排名曲线情况

采用SPSS 17.0分析预测模块中的专家建模器,选择所有模型。专家建模器可以根据模型的具体情况自动参数设定并拟合适合的模型。得到的简单季节性预测模型见图2。^⑨

图2显示,拟合值与观测值拟合的情况较好,且几乎全部落在置信区间置信上限和置信下限内。而从表2也可以看出拟合的平均绝对误差(Mean Absolute Error,简称MAE)为0.589,最大绝对误差(Maximum Absolute Error,简称MaxAE)为2.938,对于模型统计量sig为0.009,即在5%水平上显著,即“客房模型”可以很好地预测未来顾客评论在“客房”变量方面的排名情况。

根据曲线导出的预测值(表3),可以粗略

估计“客房”变量在未来的顾客评论中关注指数会呈下降趋势。

同理,对其他变量进行了预测分析,在符合模型统计量要求前提下,将变量在2012年1月至6月的预测指数前10名情况进行了汇总,得到的预测情况如表4所示。

从预测的情况可以看出,“总体服务”依然是样本酒店的顾客对其服务质量的关注点。而“方便性”、“交通”、“酒店环境”在2012年4月后将超越“客房”成为主要关注内容,而“客房”作为酒店提供给顾客的主要服务设施,顾客对其的关注在2012年上半年将基本维持在3~5名,“早餐”和“硬件设施”关注的排名也将明显偏后。

表2 “客房”变量的模型统计量

模型	模型拟合统计量				Ljung-Box Q (18)		
	平稳的R方	R方	MAE	MaxAE	统计量	DF	Sig
客房	0.759	0.396	0.589	2.938	14.999	16	0.009

表3 预测模型未来预测值情况

	2012.1	2012.2	2012.3	2012.4	2012.5	2012.6
预测指数	2.39	2.39	2.72	3.27	3.39	4.06
置信上限	4.12	4.20	4.61	5.68	5.41	6.14
置信下限	0.66	0.58	0.84	1.77	1.37	1.97

表4 2012年1~6月份顾客评论的关注面预测情况

	2012.1	2012.2	2012.3	2012.4	2012.5	2012.6
1	总体服务	方便性	总体服务	总体服务	总体服务	总体服务
2	客房	总体服务	方便性	方便性	方便性	交通
3	硬件设施	客房	客房	交通	交通	方便性
4	方便性	交通	酒店环境	客房	酒店环境	硬件设施
5	交通	硬件设施	交通	酒店环境	客房	客房
6	酒店环境	早餐	早餐	硬件设施	硬件设施	酒店环境
7	早餐	酒店位置	价格	价格	周边环境	价格
8	前台	酒店环境	硬件设施	周边环境	价格	酒店位置
9	价格	价格	周边环境	干净度	早餐	周边环境
10	舒适度	周边环境	酒店位置	早餐	推荐情况	干净度

2. 顾客评论情感分析与预测

在电子商务化的时代里，在线旅游预订渐渐成为顾客进行旅游产品交易的主流方式之一。大部分用户搜索到目标产品后，除了关注产品本身属性外，还会浏览用户评论等产品相关信息。其中41.1%的网民在购买每个产品前都看用户评论，26%的用户购买大多数产品前都会事先浏览相关评论，而负面评论对旅游预订意向具有显著影响^[24]。如果酒店在线评论中存在过多的消极情绪，可能会传递给正在预订的顾客，甚至影响到酒店的忠实顾客，并导致顾客流失。因此，随时监控与预测酒店在线评论的情感变化具有重要的实践意义。

(1) 顾客评论情感分析

情感分析是指通过自动分析某种商品评论的文本内容，发现消费者对该商品评论信息的褒贬态度和意见^[32]。

在本研究中，通过SPSS 19.0统计分析软件专家建模器建立从2009年1月1日至2011年12月31日的月份情感趋势曲线模型（图3和图4，情感趋势曲线分为积极情感分值曲线和消极情感分值曲线，分值均采用百分制），结果显示，2009年1月1日至2010年12月31日期间，顾客评论的积极情感分值曲线有较大的波动，而2011年后虽然在2011年9月出现一定的下降，但总体开始呈波动式缓慢上升趋势。而消极情感分值曲线则略显平稳，基本维持在15分以内，可见顾客评论中的消极情感表现较为稳定。

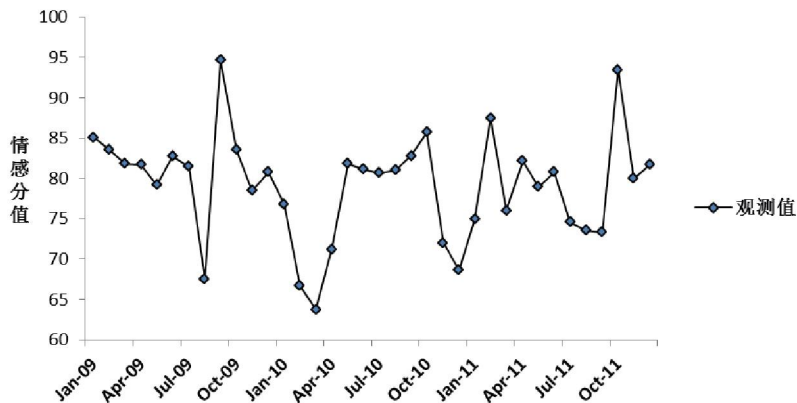


图3 积极情感评论的分值曲线情况

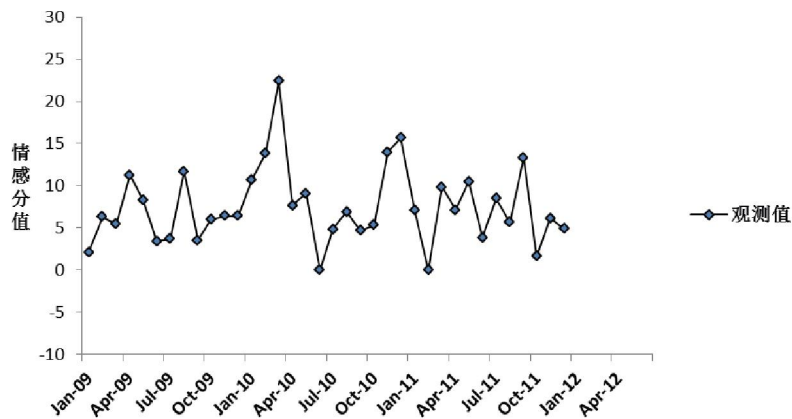


图4 消极情感评论的分值曲线情况

(2) 顾客评论情感曲线的预测

根据SPSS 19.0统计分析软件的预测功能，笔者对样本酒店顾客评论的积极情感分值曲线和消极情感分值曲线进行了预测，如图5和图6。

在积极情感和消极情感趋势图中，拟合曲线与实际曲线拟合的情况较好，且过去拟合曲线和未来预测曲线均落在置信区间内。而从表5可以看出，拟合的平均误差（积极情感：4.482；

消极情感: 3.255) 和最大误差 (积极情感: 14.140; 消极情感: 9.562) 均较理想, 在5%水平上均具有显著性, 综合认为预测效果较为理想。

根据导出的预测模型的未来预测值情况, 如

表6, 积极情感曲线分值在2012年3月将会有较大的落差, 而之后将维持在80分以上, 消极情感曲线分值从2012年1月开始至5月都将维持在较高水平, 尤其是2012年3月将达到12.23分, 而6月的消极情感曲线分值将大幅下降至2.09分。

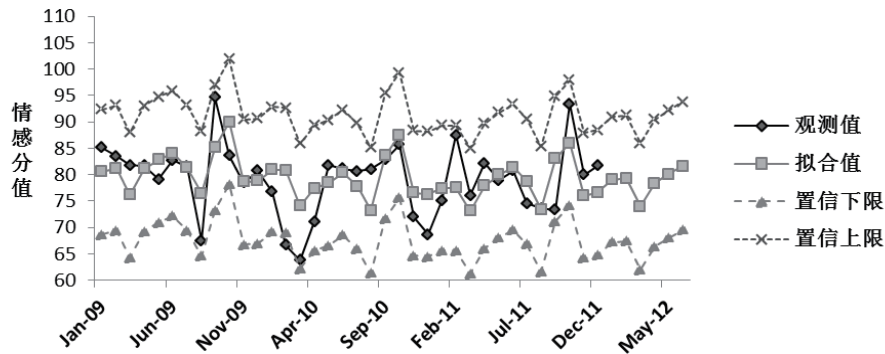


图5 积极情感评论的分值预测曲线情况

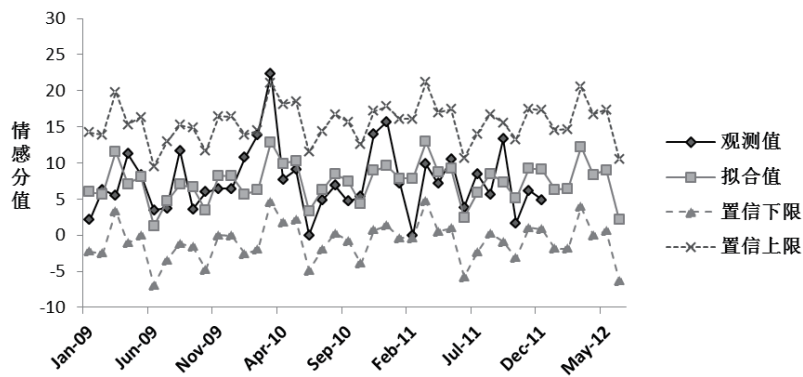


图6 消极情感评论的分值预测曲线情况

表5 模型统计量

积极情感曲线	模型	模型拟合统计量				Ljung-Box Q (18)		
		平稳的R方	R方	MAE	MaxAE	统计量	DF	Sig
	数值	0.685	0.270	4.482	14.140	46.749	16	0.000
消极情感曲线	模型	模型拟合统计量				Ljung-Box Q (18)		
		平稳的R方	R方	MAE	MaxAE	统计量	DF	Sig
	数值	0.711	0.253	3.255	9.562	46.389	16	0.000

表6 预测模型未来预测值情况

时间		2012.1	2012.2	2012.3	2012.4	2012.5	2012.6
积极情感曲线	预测值	79.06	79.33	73.98	78.42	80.07	81.64
	置信上限	90.92	91.25	85.96	90.45	92.17	93.80
	置信下限	67.20	67.41	62.00	66.38	67.97	69.49
消极情感曲线	预测值	6.32	6.40	12.23	8.36	8.98	2.09
	置信上限	14.54	14.66	20.54	16.71	17.36	10.52
	置信下限	-1.90	-1.86	3.93	0.02	0.59	-6.33

五、结论与讨论

1. 结论

本研究基于文本挖掘技术对广州市高档商务型酒店商务顾客的在线评论进行词频分析、情感分析与时间序列预测,得到如下主要结论:

第一,从词频分析的情况来看,近年来顾客评论的主要关注面是“总体服务”、“方便性”、“客房”、“交通”等,其中“总体服务”和“客房”是最主要的关注点,同时顾客评论的关注面呈现从技术质量向功能质量转移的趋势,如“方便性”、“交通”等问题的关注程度越来越高。

第二,从情感分析的情况来看,近年来顾客评论的情绪心理处于以积极情绪为主且稳步提升的状态,商务顾客对广州市高档商务型酒店的服务质量较为认可。

第三,从时间序列分析的情况来看,一方面,2012年上半年商务顾客对广州市高档商务型酒店服务质量的关注程度从高到低依次为“总体服务”、“方便性”、“酒店环境”、“交通”和“客房”,与之前提出的顾客将越来越重视功能质量的结论相一致。另一方面,2012年上半年广州市高档商务型酒店的顾客评论中会有较多带有消极情感的评论,虽然其具有积极情感的评论的百分比仍维持在70%以上,但顾客评论消极情绪的影响力不可忽视。

由此可见,顾客最为关注的是酒店的总体服务,广州市高档商务型酒店应高度重视酒店的全面质量管理,并重点提升客房服务和其他功能质量。

2. 建议

根据以上分析,笔者尝试性地对广州市高档商务型酒店的发展提出以下建议:

第一,酒店应认识到顾客评论的重要性,将其纳入服务质量管理评价体系与客户关系管理中,并将其作为服务质量管理的策略依据。第二,酒店应加强服务质量管理,特别是功能质量管理。第三,酒店应对酒店在线评论进行有效管理,一方面对参与评论的顾客进行身份认证,防止恶意评论的出现,另一方面,对消极评论应尽快回复,诚意感谢、解释或致歉,安抚不满意的顾客,并提出有效的解决措施;对做出积极评论

的顾客也要表示感谢与支持。通过各种手段的有机结合,酒店应让其评论形成更为良好的氛围,形成酒店在网络化时代下的良好电子口碑(Electronic Word-of-Mouth)。

3. 讨论

从研究方法上看,本研究弥补了酒店在线评论研究中批量化处理技术应用的缺失,从实践上为商务型酒店的服务质量管理提供了一定的参考,但依然存在一定的局限性。一方面,中文文本挖掘技术并未非常成熟,其挖掘结果的正确率还有待进一步提高;另一方面,本研究中只采集了36个月份的评论数据,虽然已符合研究采样的数据要求,但对时间序列分析的稳定性会有一定的影响,理论上讲,数据越多预测越稳定,因此所得到的预测值依然有待进一步验证。

此外,本文的后续研究可以从以下几方面进行:从研究对象类型来看,本研究可以延伸到不同类型酒店如经济型酒店或度假型酒店或者进行不同类型的酒店服务质量评价的对比分析;从研究方法上看,对于酒店顾客评论中文文本的文本挖掘,是否有更适合的挖掘技术以实现更准确的挖掘效果,也是后续研究中可以进一步深入探讨的重要内容;从研究的地域范围上看,国外商务型酒店顾客在线评论的文本挖掘研究是非常有价值的选题,如能对此进行国内外的对比分析,意义更为重大。

注释:

- ① <http://www.tripadvisor.com/>。
- ② 到到网:<http://www.daodao.com/>。
- ③ 携程网:<http://www.ctrip.com/>。
- ④ 去哪儿网:<http://www.qunar.com/>。
- ⑤ ALEXA 网站:<http://alexa.chinaz.com/>。
- ⑥ 关于 Alexa 网站排名的分析 <http://www.henan100.com/life/2011/63170.shtml>。
- ⑦ 蒋丁新. 饭店管理(第2版)[M]. 北京:高等教育出版社,2004。
- ⑧ 2010年《中国商务人士旅行调查研究报告》:<http://mfiles.sohu.com/cul/2010shangwulvxing.pdf>。
- ⑨ 观测值指的是实际获取的数据;拟合值指的是通过建模后形成的与实际数据比较接近的数据;预测值指的是依靠拟合曲线获取的代表未来的数据;置信下限和置信上限之间的区域为置信区间,即属于可信范围。下文亦按此说法说明。

参考文献:

- [1] Peter Kenzelmann. *Kundenbindung* [M]. Berlin : Cornelsen Verlag Skriptor GmbH & Co KG, 2008.
- [2] Christian G. Service quality model and its marketing implications [J]. *European Journal of Marketing*, 1984, (18): 36~44.
- [3] Akbaba A. Measuring service quality in the hotel industry : A study in a business hotel in Turkey [J]. *Hospitality Management*, 2006, 25 (2): 170~192.
- [4] Bitner M J, Booms B H, Tetreault M S. The service encounter : Diagnosing favorable and unfavorable incidents [J]. *Journal of Marketing*, 1990, 54 (1): 71~84.
- [5] Parasuraman A, Zeithaml V A, Berry L L. Servqual : A Multiple-item scale for measuring consumer perceptions of service quality [J]. *Journal of Retailing*, 1988, (64): 12~40.
- [6] Cronin J, Taylor S A. Measuring service quality : A reexamination and extension [J]. *Journal of Marketing*, 2009, 56 (7): 55~67.
- [7] 汪纯孝, 岑成德等. 服务性企业整体质量管理 [M]. 广州: 中山大学出版社, 1999.
- [8] 范秀成. 服务质量管理——交互过程与交互质量 [J]. 天津南开管理评论, 1999, (1): 8~23.
- [9] 兰琦. B2C 电子商务服务质量评价影响因素研究 [D]. 成都: 电子科技大学, 2010.
- [10] 张超鹤. 我国第三方物流企业服务质量评价及改进研究 [D]. 西安: 西安科技大学, 2011.
- [11] 高秀军. 基于顾客满意度视角的医院服务质量评价 [D]. 昆明: 云南财经大学, 2011.
- [12] 郝建南. 图书馆内部服务质量测度的实证研究 [J]. 图书馆论坛, 2011, (4): 107~110.
- [13] 戚牧. SERVQUAL 模型对 C2C 电子商务服务质量评价的实证研究 [J]. 中国商贸, 2011, (3): 110~111.
- [14] Callan R J. The critical incident technique in hospitality research : An illustration from the UK lodge sector [J]. *Tourism Management*, 1998, 19 (1): 93~98.
- [15] Gilbert D C, Morris L. The usefulness of critical incident technique analysis in isolating travel satisfaction [J]. *International Journal of Contemporary Hospitality Management*, 1995, 7 (4): 5~7.
- [16] 王文君, 韩德昌. 基于差距模型的饭店服务质量研究 [J]. 旅游学刊, 2008, (11): 19~25.
- [17] 苏志平, 顾平, 武如飞. 基于 SERVQUAL 的饭店服务质量实证研究 [J]. 商业经济, 2010, (8): 54~57.
- [18] 温碧燕. 西部地区旅游饭店业服务质量评价与研究——兼与东部沿海地区旅游饭店业比较 [J]. 经济地理, 2006, (9): 871~883.
- [19] Woo Gon Kim, Xiao jing Ma, Dong Jin Kim. Determinants of Chinese hotel customers' e-satisfaction and purchase intentions [J]. *Tourism Management*, 2006, (27): 890~900.
- [20] 朱峰, 吕镇. 国内游客对饭店服务质量评论的文本分析——以 e 龙网的网友评论为例 [J]. 旅游学刊, 2006, (5): 86~90.
- [21] Vermeulen Ivar E, Daphne Seegers. Tried and tested : The impact of online hotel reviews on consumer consideration [J]. *Tourism Management*, 2009, (30): 123~127.
- [22] Mahony M P, Smyth A B. Classification-based review recommender [J]. *Knowledge-Based Systems*, 2010, (23): 329~329.
- [23] 熊伟, 许俊华. 基于内容分析法的我国经济型酒店服务质量评价研究 [J]. 北京第二外国语学院学报, 2010, (11): 57~67.
- [24] 郭潇. 在线评论对旅游预订意向影响的实证分析 [D]. 广州: 华南理工大学, 2010.
- [25] Ronen Feldman, James Sanger. *The Text Mining Handbook* [M]. UK : Cambridge University, 2007.
- [26] 郑文英. 旅行目的地中文评论的情感分析研究 [D]. 哈尔滨: 哈尔滨工业大学, 2010.
- [27] Bing Pan, Daniel Fesenmaier R. Semantics of online tourism and travel information search on the internet : A preliminary study [J]. *Information & Communication Technologies in Tourism*, 2002 : 8~10.
- [28] 曹彬. 互联网上旅游评论的情感分析及其有用性研究 [D]. 哈尔滨: 哈尔滨工业大学, 2008.
- [29] 刘国果. 广州星级酒店空间分布特征研究 [J]. 中山大学研究生学刊, 2011, (32): 103~113.
- [30] 钟静. 高星级商务顾客感知服务质量研究 [J]. 江苏商论, 2010, (11): 26~28.
- [31] 张文彤. SPSS 11.0 统计分析教程 (高级篇) [M]. 北京: 希望电子出版社, 2002.
- [32] 张紫琼, 叶强, 李一军. 互联网商品评论情感分析研究综述 [J]. 管理科学学报, 2010, (13): 85~96.