What tweets and retweets on twitter can tell for the restaurant industry:

A big-data approach:

by

Xi Wang

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Hospitality Management

Program of Study Committee:
Liang (Rebecca) Tang, Major Professor
Eric D. Olson
Tianshu Zheng
Gang Han
Chunhui Xiang

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2020

Copyright © Xi Wang, 2020. All rights reserved.

المنارة للاستشارات

www.manaraa.com

ProQuest Number: 27834685

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 27834685

Published by ProQuest LLC (2020). Copyright of the Dissertation is held by the Author.

All Rights Reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

المنارة للاستشارات

www.manaraa.com

DEDICATION

This dissertation is dedicated to the fleet of time.

المنارة للاستشارات

www.manaraa.com

TABLE OF CONTENTS

المنارة للاستشارات

www.manaraa.com

LIST OF FIGURES

المنارة للاستشارات

www.manaraa.com

LIST OF TABLES

Page

www.manaraa.com

NOMENCLATURE

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
DMO	Destination Management Organization
ELM	Elaboration Likelihood Model
EmoLex	Emotion Lexicon
IDW	Inverse Distance Weighting
LIWC	Linguistic Inquiry and Word Count
LSM	Language Style Matching
SEM	Structural Equation Modeling
TAM	Technology Acceptance Model
UGC	User-Generated Content
VIF	Variance Inflation Factor
WOM	Word-of-Mouth

المنارة للاستشارات

www.manaraa.com

ACKNOWLEDGMENTS

I would like to thank my committee chair, Liang (Rebecca) Tang, and my committee members, Tianshu Zheng, Eric D. Olson, Gang Han, and Chunhui Xiang for their guidance and support throughout the course of this research.

In addition, I would like to thank my family, friends, colleagues, the department faculty and staff for their continued support and encouragement. I want to also offer my appreciation to Chenyu Zhao, Yitian Li, Zihao Han, Huanyue Wang, Yuxiang Xu, Yunong Liu, Dongbo Xie for making my time at Iowa State University a wonderful experience.

المنارة للاستشارات
www.manaraa.com

ABSTRACT

In the Internet age, the sheer volume of information can be generated and disseminated through online user-generated content (UGC). Within the context of Twitter, the retweeting function is one of the key mechanisms, which enables the information diffusion process among users in the social network. Stimulated by this concern, the purpose of the current study was to investigate the effects of textual content including the sentiments, emotions, and language style matching (LSM) of Twitter, a series of statistical analyses are conducted to the Twitter dataset with around one million pieces of customer tweet information. The results indicated that sentiments, emotions, and LSM have significant influences on customer retweeting behavior. Besides, significant differences were identified of both sentiments and emotions based on both six periods of the timeline analysis and the geographic distance at the city level, state level, and nationwide level. Discussions and implications interpreted the significance of the most valuable findings and suggested some important insights to both academia and industry.

Keywords: Twitter, retweeting, restaurant, emotion, language style matching

www.manaraa.com

CHAPTER 1. INTRODUCTION

With the innovation of information technology, the Internet has widely interacted in our daily life from many aspects. Benefited from the Internet, social media like other Internet-based platforms, has experienced explosive growth since the beginning of the 21st century. According to the global digital report initiated by Kemp (2019, p. 82), there were 4.4 billion people using the Internet around the world, in which, more than 80% of them were social media users in the year of 2019. People have already adapted to a new online lifestyle. For example, people like to stay in touch with each other and customers prefer to search for useful information before choosing a restaurant through the online social media (Sevin, 2013). And businesses also change their focus and make efforts in the digital marketing through a variety of social media platforms so as to increase the penetration rate in the marketplace (Minazzi, 2015, p. 80).

In the hospitality industry, literature directly established based on the information from Twitter has not been extensive, because many specific aspects in the hospitality industry has already created their own online platforms like www.booking.com or www.hotels.com for lodging, www.yelp.com for foodservice, and www.tripadvisor.com and www.expedia.com from tourism as a general. Therefore, these websites or platforms provide valuable and rich content for studying and analyzing online user-generated content (UGC) and even customer behavior on the Internet as a general. Furthermore, due to the technical difficulties, many researchers without the sufficient information technology background were discouraged to develop topics or conduct data analysis based on large-scale data on social media.

As one of the main players in social media, Twitter with more than 320 million

المنارة للاستشارات
www.manaraa.com

monthly active users, is ranked the second position in the micro-blog businesses in the world and makes essential influences on people from many aspects such as communication style, shopping behavior, and even business model (Kemp, 2019, p. 183). Besides the huge user base, Twitter on the one hand, has been indicated as one of the most heavily engaged platforms by people, and at the same time, it has also become an important marketing channel to reach its audiences for different kinds of businesses and create competitive advantages in the market (Hay, 2010; Philander & Zhong, 2016). In the meantime, UGC in Twitter has been long regarded as a meaningful source which can be used to reveal very important information related to customer behavior when doing research (Misopoulos et al., 2014). Filtering with keywords related to hospitality, information from Twitter can be precisely extracted and selected to do the study targeting to the hospitality industry as well.

Previous studies in the hospitality and tourism field analyzing the Twitter data have been not as abundant as those in the general marketing or communication disciplines. The primary areas or topics examined in the existing studies of the hospitality and tourism field included customer behavior which focuses on purchasing activities associated with individual or group activities, UGC which focuses on analyzing information provided by customers, destination management organization (DMO) which focuses on investigating approaches and strategies of tourist destinations, latent topic which focuses on developing in-depth research from complex information computed by algorithms, and electronic word-of-mouth (eWOM) which focuses on analyzing buzz marketing and reputation of the business. And these previous studies are summarized in Table 1.

المنارة للاستشارات

www.manaraa.com

Table 1. Summary of Twitter studies in hospitality and tourism industries

Topic and context	Main method	Reference
Customer behavior		
Tourism	Machine learning with support vector machine	Shimada, Inoue, & Endo (2012)
Restaurant	Correlation	Abbar, Mejova, & Weber (2015)
Hotel	Sentiment analysis	Philander & Zhong (2016); Phillips-Wren & Hoskisson (2014)
Hotel	Ordinary least squares (OLS)	De Rosario, Rodríguez, & Pérez (2013)
User-generated content (UGC)		
Tourism	Sentiment analysis, and self-organizing map (SOM)	Claster, Cooper, & Sallis (2010); Claster et al. (2013)
Tourism	SNA - mapping	Park, Ok, & Chae (2016)
Restaurant	Sentiment analysis	Park, Jang, & Ok (2016)
Airline	Sentiment analysis	Misopoulos et al. (2014); Mostafa (2013)
Airline	Sentiment analysis and sentiment topic matching	Adeborna & Siau (2014)
Cruise	SNA - mapping	Park, Ok, & Chae (2016)
Destination management organization (DMO)		
Tourism	Descriptive analysis	Hay (2010)
Tourism	Social network analysis (SNA)	Antoniadis, Zafiropoulos, & Vrana (2015); Bokunewicz & Shulman (2017)
Tourism	Content analysis and semistructured interview	Hays, Page, & Buhalis (2013)
Latent topic		
Tourism	Spatial analysis and textual analysis	Brandt, Bendler, & Neumann (2017)
Restaurant	Content analysis	Vidal et al. (2015)
Restaurant	RCRP-LDA analysis (Recurrent Chinese Restaurant Process & Latent Dirichlet allocation)	Diao & Jiang (2014)
Electronic word-of-mouth (eWOM)		
Tourism	Ordinary least squares (OLS)	Sotiriadis & Van Zyl (2013)
Restaurant	Logistic regression	Jiang & Erdem (2017)
Restaurant	Real-time sentiment analysis	Jansen et al. (2009)
Airline	Partial least squares-structural equation model (PLS-SEM)	Vo et al. (2017)

المنارة للاستشارات
www.manaraa.com

The current study integrally considered both customer behavior and online UGC as two primary components. Referring to Table 1, although previous studies discussed UGC and conducted sentiment analysis based on these contents, they even hardly had an in-depth analysis of the individual emotion level. For example, Park, Jang, and Ok (2016) investigated the sentiments in the restaurant context, but it only regarded sentiments as a whole score and compared the restaurants among different counties.

However, taking a further step based on previous studies, the current study took not only the general two sentiments (i.e., positive vs. negative), but also the eight emotional dimensions (i.e., anger, anticipation, disgust, fear, joy, sadness, surprise) into consideration. Because based on the theory of the Elaboration Likelihood Model (ELM), although people nowadays normally would not like to conduct a precise assessment on online information and just follow the peripheral route of persuasion, emotions expressed along with textual information of the tweet content can influence the attitude of people. And the same with customer behavior in many businesses, once there is a change in the attitudes and perceptions, there would be corresponding chain effects on customer behavior intention. Therefore, both customer sentiment and emotion were two critical components when measuring the online UGC. Due to the importance of sentiments and emotions considered in the current study, two more analysis including timeline analysis and geostatistical analysis were also applied as the additional analysis focusing on them.

In addition, language style matching (LSM) was another perspective investigated in the current study. LSM mainly measures language structure and synchronized verbal behavior (verbal mimicry) of textual information generated by people. LSM was important in the textual analysis because the verbal mimicry normally appears with social

المنارة للاستشارات
www.manaraa.com

behavior, and it can be used to reflect the matching of identical behaviors among people (Gonzales, Hancock, &Pennebaker, 2010). And through an examination of nine functional words including personal pronouns, impersonal pronouns, articles, conjunctions, prepositions, auxiliary verbs, high-frequency adverbs, negations, quantifiers, LSM was used as an indicator of cohesiveness and performance of the textual content provided among people (Gonzales, Hancock, & Pennebaker, 2010). Therefore, along with the sentiments and emotions, LSM in the current study was also taken into consideration to detect the matching behavior of the tweets posted by customers.

Within the context of the hospitality industry, previous literature has rarely involved with the exploration of retweeting behavior of customers. As to Twitter, retweeting function is designed as one of the key mechanisms in Twitter, and it is regarded as an important bond that enables the process of information diffusion among users in the Twitter social network (Suh et al., 2010). When people retweet, they can easily mention and cite some viewpoints from original tweets; meanwhile, people can display their public agreement or disagreement with some popular topics.

On the other hand, through retweeting, not only the textual information, but also other information embedded with the UGC including sentiments, emotions, and functional words can be transferred and expressed by potential readers along with the information communication process (Boyd, Golder & Lotan 2010). And based on the ELM theory, emotions embedded in tweets can determine people's behavioral intentions according to the peripheral route of persuasion (Petty & Cacioppo, 1986, p125). Therefore, in the current study, customers' retweeting behavior which is measured by the retweet number is considered as the important variable of interests which it is worthy of

المنارة للاستشارات
www.manaraa.com

exploring in the hospitality industry.

Based on the discussion above, the primary purpose of the current study was to explore some determinants in Twitter from both customer profile and UGC information as well as customer retweeting behavior. Specifically, through a series of text mining processes, estimation procedures, computing algorithms, and regression analyses, the current study was to: 1) investigate the effects of two sentiments (i.e., positive vs. negative) embedded in a certain tweet on the customer retweeting behavior; 2) examine the effects of eight emotional dimensions embedded in a certain tweet (i.e., anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) on the customer retweeting behavior; 3) evaluate the effects of the total LSM score of a certain tweet on customer retweeting behavior. Besides, the current study also took a further step to examine the customer sentiments and emotions from the geostatistical and timeline perspectives as well.

المنارة للاستشارات
www.manaraa.com

CHAPTER 2. LITERATURE REVIEW

2.1. Twitter and its characteristics

Since the establishment of Twitter in 2006, organizations have gradually expanded their businesses and increased their marketing activities through this new online channel. At the same time, more and more customers have accepted this new communication approach to post and share information online (Bokunewicz & Shulman, 2017; Hay, 2010). As to the hospitality industry, social media such as Twitter and Facebook has provided a new business model to the DMO which reshapes the management strategies from many aspects, including services, marketing, networking, and knowledge management (Zeng & Gerritsen, 2014). Due to the openness feature and popularity among customers, the great marketing potential of Twitter has also been discovered by the hospitality businesses (Hay, 2010), and it even can be treated as the primary channel for marketing and promotions (Park, Ok, & Chae, 2016).

Like some general characteristics of social media, Twitter allows users creating and posting information with a variety of formats such as text and picture (Thelwall, Buckley, & Paltoglou, 2011), and provides a network platform that users can make comments and communicate with each other in a very easy way (Hay, 2010). However, due to its business segment targeting to the micro-blog products, several unique features can help fulfill the deficiency inherited from other social media platforms such as Facebook. First, the length of the post is the primary indicator that makes Twitter distinctive from other platforms. With the micro-blog feature, the length of each tweet (Twitter post) is limited to 280 characters. However, due to the limitation of the characters, a quick-broadcasting feature is incisively utilized by users (Weng et al.,

www.manaraa.com

2010). For example, Sakaki, Okazaki, and Matsuo (2010) indicated that there is a significant increase in real-time interactions between tweets and an important event such as earthquake. The second unique feature is retweet. The retweeting function allows users to make minor modifications, and then forward and post the original tweet again very easily. Taking advantage of the retweeting function, it not only ensures an extremely rapid information spread, but also significantly increases Twitter’s broadcasting efficiency (Thelwall, Buckley, & Paltoglou, 2011). The third feature is the hashtag. When posting information on Twitter, users can include some hashtags of keywords so as to post tweets under certain topics. By using the hashtags, it provides a convenient way for users to find a piece of tweet information and also enables the communication and interaction among users within a certain topic (Thelwall, Buckley, & Paltoglou, 2011).

Besides, the mobility of Twitter is also another advantage preferred by the majority of its users. Indicated by the global digital report, among those 3.5 billion active social media users, more than 98% of them are mobile social media users (Kemp, 2019, p. 76), that is, comparing with using computers or laptops, accessing to the social media from mobile devices has already dominated the market. And just incorporating with the feature of mobility, Twitter launched its mobile application on many different platforms at a very early stage, which brings it to the leading position compared with other brand followers (Java et al., 2007).

Table 2 summarizes and compares the user’s post functions among the primary social media platforms. Twitter does not include functions of directly evaluating a certain business, which is different from the customer review platforms such as TripAdvisor, Booking.com, and Yelp. With the unique features discussed above, Twitter is regarded by

المنارة للاستشارات
www.manaraa.com

the young generation as one of the most popular social media platforms in the world to share with their opinions and show the attitudes (Sotiriadis & Van Zyl, 2013).

Table 2. Summary of posting functions of social medial platforms

Feature	Twitter	Facebook	TripAdvisor	Booking.com	Yelp
Type	Micro-blog	Blog	Review site	Review site	Review site
Character limit	280	63,206	200	N/A	5,000
Context	Social networking	Social networking	Tourism	Hotel	Restaurant
Mention user	Yes	Yes	No	No	No
Link	Yes	Yes	No	No	No
Hashtag	Yes	Yes	No	No	No
Post picture	Yes	Yes	Yes	No	Yes
Post Date	Yes	Yes	Yes	Yes	Yes
Check-in	No	No	Yes	Yes	Yes
Location	Yes	No	Yes	Yes	Yes
Highlight	No	No	Yes	Yes	No
Advantage and disadvantage	No	No	No	Yes	No
Useful/like	Yes	Yes	Yes	Yes	Yes
Overall rating	No	No	Yes	Yes	Yes
Detail rating	No	No	Yes	No	No
User profile	Yes	Yes	Yes	No	Yes
User interaction	Yes	Yes	Yes	No	Yes

2.2. Twitter in hospitality

Although the micro-blog function of Twitter only allows users to post a tweet under 280 characters and below, it does not vanish the enthusiasm of people to communicate with each other at all (Park, Ok, & Chae, 2016). And just because of the convenient tweet and retweet functions, a large volume of online UGC can be collected which can bring valuable information to both its users and marketers (Sotiriadis & Van Zyl, 2013). In the hospitality industry, UGC has become a prevalent topic and explored by previous studies in recent years. According to the summary of Philander and Zhong (2016), UGC from Twitter has a significant influence on the hospitality industry from a variety of research topics including branding, marketing, finance, and customer

المنارة للاستشارات
www.manaraa.com

satisfaction. Analyzing the UGC not only helps businesses in the hospitality industry to know their customers' attitudes toward a product or a brand in an intuitionistic way (Minazzi, 2015, p. 34), but also contributes to understand their customer purchasing behavior and decision-making processes (Sotiriadis & Van Zyl, 2013). Therefore, the importance of studying both Twitter and its UGC is self-evident in the hospitality industry.

Previous literature related to the topic of Twitter in social science and business administration fields were usually discussed from five aspects: a user's attitude towards Twitter (e.g. Java et al., 2007), tweeting and retweeting behavior (e.g. Boyd, Golder, & Lotan, 2010), textual mining of tweet (Agarwal et al., 2011), network analysis (e.g. Kulshrestha et al., 2012), and predictive analytics with Twitter information (Bollen, Mao, & Zeng, 2011). However, when targeting the hospitality industry, literature directly focusing on Twitter has not been as adequate as other disciplines. Because on the one hand, besides Twitter there has been other specific social media platforms in the hospitality industry like the Booking.com which focuses more on the hotel businesses, and Yelp platform which focuses more on the restaurant business. The studies developed based on these identical social media platforms also can access to a relatively large volume of UGC and achieve the research purpose related to the textual mining (e.g. De Albornoz et al., 2011; Wang, Tang, & Kim, 2019). On the other hand, due to the technical difficulty, it would be challenging when doing the web scripting and collecting from the Twitter platform comparing to other simple designed websites such as Yelp (Alaimo, 2018). According to the previous studies, the only way to obtain the data of Twitter was to program based on the Twitter Application Programming Interface (API)

المنارة للاستشارات
www.manaraa.com

(Agarwal et al., 2011; Bifet & Frank, 2010). However, considering that the technical challenge is one of the considerations of researchers in any field, it has increased the barrier of doing studies of Twitter at the data collecting stage.

2.3. Retweeting behavior

According to Table 1 in the introduction section, although studies somehow have discussed customer behavior by using twitter information, few of them have investigated customer retweeting behavior and hardly examined the relationship between tweeting and retweeting in the hospitality industry.

As mentioned above, Twitter provides the micro-blog service that allows various formats of information and enables a convenient way to communicate with each other (Hay, 2010; Thelwall, Buckley, & Paltoglou, 2011). Retweeting or retweeting behavior is one of the key mechanisms in Twitter, which enables the process of information diffusion among users in the Twitter social network (Suh et al., 2010). Unlike making comments or sharing of Facebook updates, retweeting is a prototypical way to repost the original message and/or include an individual's own contents (Boyd, Golder, & Lotan, 2010). Preceding with "RT" and original author name with "@" syntax, the retweeting behavior can be easily processed with one-click function of Twitter on both desktop software and mobile app. And an example of the retweeting format was provided as follows.

A: I like Starbucks!

B: RT @A: I like Starbucks!

People retweeting a certain tweet can have several reasons. Discussed by Suh et al. (2010), mentioning is one of the particular cases when people retweet. For example, when people find any interesting topic and want to share with others, the retweeting function can directly include the original information which lets other people easily know

المنارة للاستشارات

www.manaraa.com

the information source. Retweeting behavior also can invite some new people (users of original tweets) to engage in a particular thread by mentioning them, however, without directly talking to them (Boyd, Golder & Lotan 2010). At the same time, deemed by Stieglitz and Dang-Xuan (2012), in addition to the basic topic or information in tweets, the retweeting behavior can in the meantime bring certain values from the original tweets or users to other people. For example, people may retweet information posted by authority, mention them, comment them, and show publicly agreement or disagreement with them, which further support their personal viewpoints.

From the information diffusion perspective, the importance of retweeting behavior is self-evident. On the one hand, retweeting can be regarded as a mean of participating, and through the process of spreading tweets (information) to other people, retweeting behavior can increase the engagement between each other (Boyd, Golder & Lotan 2010). On the other hand, with the considerable engagements of tweeting and retweeting behavior, the Twitter business can not only consolidate its position in the micro-blog industry, but also maintain the competitive advantage in the market that continuously attract many kinds of business to promote and advertise through Twitter and prefer it as a primary marketing channel. (Hay, 2010; Philander & Zhong, 2016).

As mentioned in section 2.2., previous studies related to the retweeting behavior were adequate in the fields of general marketing and customer behavior. However, the literature explored the determinants on retweeting behavior contained several aspects including user profile (e.g. Luo et al., 2013), tweet language (e.g. Molyneux, 2015), tweet topic (e.g. Macskassy & Michelson, 2011), and tweet sentiments (e.g. Stieglitz & Dang-Xuan, 2012). Unfortunately, limit studies investigated those potential topics specifically

المنارة للاستشارات
www.manaraa.com

targeting to the hospitality industry. Therefore, with this concern, taking an in-depth analysis of customer retweeting behavior and examining the main determinants from both customer profile and tweet contents were the primary purposes of the current study.

2.4. Elaboration Likelihood Model (ELM)

Initiated by Petty and Caciopo (1981), the Elaboration Likelihood Model (ELM) was originally developed as a framework to explore different ways of effective persuasion in the social-psychological area. And concluded in the ELM theory, there are two major persuasion routes: the central route and peripheral route, which have different effects on the persuasive communication process and also lead to the attitude change of people. Indicated in the ELM theory, the central route of persuasion is normally associated with “a person’s careful and thoughtful consideration of the true merits of the information presented in support of an advocacy” (Petty & Cacioppo, 1986, p125); while the peripheral route of persuasion is “more likely occurred as a result of some simple cue in the persuasion context (e.g. attractive cue) that induced change without necessitating scrutiny of the true merits of information presented” (Petty & Cacioppo, 1986, p125).

According to the definition of both these two routes, one of the most fundamental distinctions between the two persuasion routes is the level of consideration when people face and compare the merits of information presented. With an example provided by Lumen (2019), if a small business owner would like to purchase a certain type of computer, he or she would follow the central route and be persuaded and influenced by the information quality and computer features such as processing speed and memory capacity. However, if some young generations would like to buy sneakers advertised by a famous athlete, he or she would follow the peripheral route because it does not require much effort in information processing. Both the central route and peripheral route can

المنارة للاستشارات
www.manaraa.com

influence people when processing persuasive information and make a difference in their attitude changes simultaneously.

Table 3. ELM literature summary since 2010

Dependent variable	Data	Method	Reference
General hospitality			
Review helpful, funny, cool	Online review	Regression	Li et al. (2017)
Sharing intention	Survey	Regression	Hur et al. (2017)
Hotel			
Purchase intention	Survey	Regression	Cheng & Loi (2014); Lwin & Phau (2012)
Booking intention	Survey	Regression	Zhao et al. (2014)
Booking intention	Survey	ANOVA	Tsao et al. (2015); Xie et al. (2011)
Customer perception	Experiment	MANOVA	Hu (2012)
Loyalty	Survey	Regression	Levy & Duverger (2010)
Service recovery	Survey	SEM	Tsao (2018)
Trustworthiness	Experiment	SEM	Pan & Chiou (2011)
Restaurant			
Perceived risk	Survey	SEM	Hussain et al. (2017)
Review helpful, funny, cool	Online review	Regression	Park & Nicolau (2015); Wang, Tang, & Kim (2019); Yin et al. (2014),
Information adoption	Survey	Regression	Salehi-Esfahani et al. (2016)
Tourism			
Purchase intention	Survey	Regression	Loda (2011)
Use intention	Survey	Regression	Ayeh, Au, & Law (2013)
Behavioral change	Survey	SEM	Chung, Han, & Koo (2013); Chung & Han (2017)
Visiting intention	Survey	SEM	Wang (2015)
Continue intention	Survey	Regression	Chung et al. (2015)
Continued usage	Survey	SEM	Kim et al. (2016)
Satisfaction	Survey	SEM	Yoo et al. (2017)
Other			
Satisfaction and Relevancy	Survey	Regression	MacDonald, Milfont, & Gavin (2016)
Behavioral intention	Survey	ANOVA	Brown, Ham, & Hughes (2010)

المنارة للاستشارات

www.manaraa.com

Table 3 summarizes the studies established based on the ELM in the hospitality and tourism industry since 2010. According to Table 3, research related to the customer behavior intention include sharing intention (e.g. Hur et al., 2017), purchase intention (e.g. Cheng & Loi, 2014; Lwin & Phau, 2012), booking intention (e.g. Tsao et al., 2015; Zhao et al., 2014), and use intention (e.g. Ayeh, Au, & Law, 2013; Kim et al., 2016) was identified as the most favored topic developed from ELM theory. Meanwhile, with an effect on the people's attitude change in the persuasive communication process, ELM also can be applied to measure customer loyalty (e.g. Levy & Duverger, 2010), satisfaction (e.g. MacDonald, Milfont, & Gavin, 2016; Yoo et al., 2017), and service (e.g. Tsao, 2018). Besides, while moving ahead with information technology, considering and solving problems with the big data techniques provides a new way to access to those problems and also be regarded as one of the new research trends in the hospitality and tourism industry. For example, taking advantage of the UGC such as customer review content and customer rating information, Wang, Tang, and Kim (2019) analyzed the customer perceived usefulness and Zhu, Yin, and He (2014) checked the customer perceived helpfulness, with the textual mining method on those thousands of review content provided by each customer online.

2.5. Sentiment and emotion

As mentioned above, ELM was selected as the theory based to explore both sentiment and emotion in the current study because it directly measures the influence process and its corresponding effects on people's perceptions and behaviors. Based on the theory of ELM, when processing the online UGC information requires thoughtful considerations, people are motivated to follow the central route process in ELM and are highly involved in elaborating the textual information (Peng et al., 2014).

المنارة للاستشارات

www.manaraa.com

However, although the central route of persuasion dominates the impact of attitude changes (Petty, Barden, & Wheeler, 2009), people nowadays would more frequently use the peripheral route and just have a roughly glance on some information and then change both attitude and behavior to a certain product or service. Because with the accelerated pace of our daily life, it has become impossible and unrealistic for people to have a careful review of every piece of information appeared anywhere around them (Petty, Barden, & Wheeler, 2009). Therefore, when analyzing the UGC, both central and peripheral routes in ELM can coincide and jointly influence people's attitudes and behaviors. As to the hospitality and tourism industry, Table 4 and Table 5 present the summary of the previous studies which conducted sentiment analysis or emotion analysis since 2010.

Revealed in these tables, research related to sentiment and emotion has been sufficiently analyzed among all areas including hotel, restaurant, and event. However, previous studies are likely to cover many aspects, compared with the content in these two tables, analyses of emotions are not as thoughtful as sentiment. First, when comparing Table 4 and Table 5, most of the studies have considered emotion as a construct which has not been measured from the specific dimensions (e.g. Brunner-Sperdin, Peters, & Strobl, 2012; Meng & Choi, 2017). And despite some research has tried to measure emotions in a precise way, only one or just a few dimensions have been analyzed and reported in previous studies (e.g. Han & Jeong, 2013; Min & Kim, 2019). However, in the current study, there were eight different kinds of emotions were detected and explored from the UGC. This part was discussed in the next section. Second, as to the research method, it was not hard to see that most of the studies selected to use the survey method

المنارة للاستشارات

www.manaraa.com

and were conducted with the structural equation modeling (SEM) analysis. However, in the current study, the text mining method by using emotion lexicon (EmoLex) analysis was used to detect both sentiment and emotions based on the Twitter dataset with the big data technique. It contributed to investigating textual information with an innovative and advanced research method and having significant contributions from both academic and practical perspectives in the hospitality and tourism industry.

Referring to Table 4, although emotions have been widely examined in previous studies on UGC like Tweets, most previous studies only stopped at the discussion on the sentiment analysis from positive vs. negative aspects (e.g. Da Silva, Hruschka, & Hruschka, 2014). In addition, although some previous studies mentioned one or several specific emotions such as anger (e.g. Min & Kim, 2019), in the hospitality and tourism industry, there has been still a very significant research gap of integrally analyzing on the effect of every individual emotion detected from the UGC. Therefore, in order to further investigate customer emotions expressed along with textual content, an eight-dimension emotional framework from the psychology discipline proposed by Plutchik was used to help detect and measure those emotions in the current study (Wang, Tang, & Kim, 2019).

المنارة للاستشارات

www.manaraa.com

Table 4. Sentiment analysis review in hospitality and tourism since 2010

Industry and data source	Sentiment	Dependent variable	Method	Reference
General hospitality				
Twitter	Two sentiments	Emotion	Descriptive analysis	Philander & Zhong (2016)
Survey	Negative emotion	Satisfaction	ANCOVA	Wu, Mattila, & Han (2014)
Secondary data	Consumer sentiment index	Stock return/ Expenditure	Regression	Singal (2012)
Hotel				
Online review website	Two sentiments	Expectation	Other	Liu et al. (2013)
Survey	Two sentiments	Loyalty	SEM	Jani & Han (2015)
Survey	Two sentiments	Willing to disclose information	SEM	Morosan & Defranco (2015)
Survey	Negative emotion	Behavioral intention	SEM	Xie & Heung (2012)
TripAdvisor	Two sentiments	Sentiments	Descriptive analysis	Yadav & Roychoudhury (2019)
TripAdvisor	Sentiment polarity	Overall rating	Regression	Zhao, Xu, & Wang (2019)
Restaurant				
Survey	Two sentiments	Loyalty	SEM	Peng, Chen, & Hung (2017)
Survey	Two sentiments	Satisfaction	ANOVA	Miao, Mattila, & Mount (2011)
Survey	Two sentiments	Satisfaction	SEM	Jung & Yoon (2011)
Survey	Two sentiments	Satisfaction	SEM	Song & Qu (2017)
Survey	Positive emotion	Behavioral intention	SEM	Jang, Ha, Park (2012)
Survey	Negative emotion	Behavior intention	Chi-square	Kim & Jang (2014)
Survey	Negative emotion	Brand attitude	ANOVA	Hwang & Mattila (2019)
Experiment	Positive emotion	Satisfaction	ANOVA	Kim & Jang (2015)

Note: Structural equation modeling (SEM); Analysis of variance (ANOVA); Analysis of covariance (ANCOVA).

المنارة للاستشارات
www.manaraa.com

Table 5. Emotion analysis review in hospitality and tourism since 2010

Industry and data source	Emotion	Dependent variable	Method	Reference
General hospitality				
Experiment	Emotion	Consumer response	ANCOVA	Wu, Mattila, & Hanks (2015)
Survey	Emotion	Behavioral intention	Regression	Chang (2016)
Survey	Pleasant	Memory Behavioral intention	SEM	Loureiro (2014)
Survey	Pleasure/arousal	Satisfaction Approach avoidance	SEM	Lin & Worthley (2012)
Survey	Emotion	Satisfaction Quality of life	SEM	Meng & Choi (2017)
Hotel				
Airbnb	Eight emotions	Overall rating	Other	Luo & Tang (2019)
Experiment	Emotion	Emotions	ANOVA	Siamionava, Slevitch, & Tomas (2018)
Survey	Emotion	Behavioral intention	Regression	Tantanatewin & Inkarojrit (2018)
Survey	Emotion	Satisfaction	SEM	Brunner-Sperdin, Peters, & Strobl (2012)
Survey	Emotion	Satisfaction index	SEM	Deng, Yeh, & Sung (2013)
Survey	Emotion	WOM	SEM	Sukhu et al. (2019)
Restaurant				
Survey	Pleasure	Behavioral intention	SEM	Hyun, Kim, & Lee (2011)
Survey	Pleasure/arousal	Emotion Satisfaction	SEM	Ruiz-Alba et al (2019)
Survey	Emotion and sentiment	Purchase intention	MANCOVA	Kim, Youn, & Rao (2017)
Survey	Four dimensions of emotion	Satisfaction/loyalty	SEM	Han & Jeong (2013)
Survey	Anger	Voice	SEM	Min & Kim (2019)
Survey	Emotion	Willing to use	Nomological validity testing	Lu, Cai, & Gursoy (2019)
Yelp	Eight emotions	Review helpfulness	Regression	Wang, Tang, & Kim (2019)
Event				
Survey	Enjoyment	Behavioral intention	SEM	Lee, Xiong, & Hu (2012)
Survey	Pleasure/arousal	Satisfaction Loyalty	SEM	Carneiro et al. (2019)
Other				
Survey	Emotion	Experience	Regression	Torres et al. (2019)
Survey	Pleasure	WOM satisfaction	SEM	Loureiro, Almeida, Rita (2013)

Note: Word of mouth (WOM); Structural equation modeling (SEM); Analysis of variance (ANOVA); Multivariate analysis of covariance (MANCOVA).

المنارة للاستشارات
www.manaraa.com

2.6. Multi-dimensional emotional framework

In previous literature, human emotions have been widely explored from different perspectives. For example, focusing on the affective information in textual content, Francisco and Gervas (2006) developed an algorithm automatically classifying text content into three-dimensional emotion categories including pleasantness, activation, and dominance. In addition, investigated by Ekman (1992), there were six distinct emotion dimensions including joy, sadness, anger, fear, and disgust, which were regarded as basic emotions when people deal with all fundamental life-tasks. Plutchik (1994) expanded the basic emotion concept with two more components of trust and anticipation, and demonstrated a framework of eight-emotional detentions.

According to Plutchik (1994)'s emotional framework, eight emotions positioned in an emotion wheel presented in Figure 1, and were divided into four opposite pairs including sadness-joy, fear-anger, disgust-trust, and surprise-anticipation. In the meantime, all pairs of emotions can be identified from the two sentiment polar (i.e., positive and negative).

المنارة للاستشارات
www.manaraa.com

Figure 1. Eight-dimensional emotion wheel

The reasons for choosing those eight basic emotions provided in Plutchik (1994)'s framework into analysis were consisted of threefold. First, the emotional framework was a well-established psychological field with adequate theoretical background. Second, the emotional framework not only considers discrete emotions, but also evaluates them into four opposite pairs, which further enables the possibility to explore the relationship between each other. Last but not least, the emotional framework has been widely used in analyzing emotions of the text content within the social science context including analysis settings on the Internet, such as general online reviews (e.g., Felbermayr and Nanopoulos, 2016), blogs (e.g., Gill et al., 2008), and micro-blog (e.g. Kim, Bak, & Oh,

المنارة للاستشارات
www.manaraa.com

2012). Meanwhile, it was also introduced in the studies in the hospitality industry, analyzing customer reviews and inspecting customer behaviors (Wang, Tang, & Kim, 2019). Therefore, with these concerns, ten hypotheses were proposed as follows.

H1. Positive sentiment embedded in tweets has a positive impact on customers' retweeting behavior.

H2. Negative sentiment embedded in tweets has a positive impact on customers' retweeting behavior.

H3. Anger emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H4. Anticipation emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H5. Disgust emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H6. Fear emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H7. Joy emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H8. Sadness emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H9. Surprise emotion embedded in tweets has a positive impact on customers' retweeting behavior.

H10. Trust emotion embedded in tweets has a positive impact on customers' retweeting behavior.

المنارة للاستشارات
www.manaraa.com

2.7. Language style matching (LSM) and functional words

In addition to sentiment and emotion, the current study was also interested in language structure and synchronized verbal behavior (verbal mimicry) of tweets posted by customers. Introduced by Gonzales, Hancock, and Pennebaker (2010), verbal mimicry normally appears when having a social behavior among people, and it reflects the matching of identical behaviors between each other. One classical example is imprinting behavior entrainment between mothers and infants. However, rather than just imitating some nonverbal behavior or movement, people can also impersonate some verbal contents in conversation or speech based on the quality of a relationship (Giles & Coupland, 1991). For example, two people who are in the same industry work in a different office building, although their daily conversation sentences and topics are different, they also tend to use some function words similar to the extent that the friends like (Ireland et al., 2011). And this kind of verbal mimicry can normally occur in the syntactic structure and even word-by-word matching (Pickering & Garrod, 2004). Therefore, targeting to the verbal mimicry calculating, an LSM algorithm based on the function words was developed by Gonzales, Hancock, and Pennebaker (2010).

Discussed by Gonzales, Hancock, and Pennebaker (2010), LSM is a type of automated measure of verbal mimicry, which can be objective, efficient, and unobtrusive when computing the verbal mimicry and functional words and predicting changes in social psychological factors of interest. With an examination of nine functional words including personal pronouns, impersonal pronouns, articles, conjunctions, prepositions, auxiliary verbs, high-frequency adverbs, negations, quantifiers, LSM can be regarded as language indicators of cohesiveness and performance for the textual content provided among people (Gonzales, Hancock, & Pennebaker, 2010). And at the same time, LSM is

المنارة للاستشارات
www.manaraa.com

proven with a high accuracy when predicting critical real-world behaviors in a different context (Ireland & Pennebaker, 2010). The nine functional word categories and some sample words in each category are presented in Table 6.

Table 6. Categories and sample words of language style matching

Category	Examples
Personal pronouns	he, she, you, we, they
Impersonal pronouns	this, it, other, who, what
Articles	a, an, the
Conjunctions	although, and, or, so, unless
Prepositions	across, below, over, since, despite
Auxiliary verbs	are, am, is, have, do
High-frequency adverbs	about, beyond, anyway, herein, maybe
Negations	never, neither, cannot, nor, nope
Quantifiers	add, both, any, lot, much

Although LSM was established based on social psychological field, it was a relatively new concept grown around 2010 and not broadly explored in the hospitality industry, among which Wang, Tang, and Kim (2019) directly applied LSM as a variable of interests and used the score calculated by the LSM algorithm to investigate the customer review helpfulness in the restaurant industry. And according to the theory of the Technology Acceptance Model (TAM), the perceptiveness of customers can finally affect customer behavior intention. Therefore, considering a rationale that customer tweets in the specific context of restaurant or dining experience should have a common group of verbal expression habit, LSM score was also taken into consideration as an another independent variable of interest to explore the influence on the customer retweeting behavior. And the hypothesis was proposed below.

H11. LSM score calculated in tweets has a positive impact on customers' retweeting behavior.

المنارة للاستشارات

www.manaraa.com

2.8. Geostatistical analysis

Defined by Goovaerts (1997), geostatistics was specialized in analysis and interpretation data with geographically attributes, and geostatistics was increasingly preferred as a method when it is necessary to capitalize the geographic and spatial relationship between neighboring observations. Introduced by Hengl (2009, p3), the spatial prediction (spatial interpolation) was the primary concern of geostatistical analysis, and with the prediction attribute, geostatistics was usually used to estimate values at those unsampled areas of interest (Goovaerts, 2000). And according to the actual measurements and semi-automated algorithms, the geostatistical analysis is different from traditional prediction approaches, and always presents the results from geostatistics through the geostatistical mapping method and provided with a visualization based on the perdition results (Hengl, 2009, p3). Considering that geographic influence and using geostatistics can contribute to the analysis of data with spatial attributes and generate information and knowledge that can be used for a more precise prediction (da Silva et al., 2012).

Within the hospitality context, although the spatial analysis and areal data have become popular in recent years, for example, several studies investigated hospitality by using the spatial regression or indicated the spatial relationship through Moran's I from many aspects such as hotel room rate (e.g. Balaguer & Pernías, 2013), hotel location (e.g. Adam & Mensah, 2014), and hotel pricing model (Tang, Kim, & Wang, 2019). However, limited research delved into the geostatistical data and investigated the continual effects around a location without a series of limitations and boundaries manually set by human (Bivand et al., 2008, p263). Therefore, with the concern from the geostatistical perspective, in addition to exploring the effects of customer textual tweet information on

www.manaraa.com

the retweeting behavior, the current study applied the geostatistical analysis and innovatively had an estimation and visualization of both sentiments and emotions on the city level, state level, and national wide level based on the US map.

المنارة للاستشارات

www.manaraa.com

CHAPTER 3. METHODOLOGY

As discussed above, the main purpose of the current study was to explore latent emotions accompanied by their tweets, investigate the difference of emotional expressions with timeline sequence, and determine the influential factors on customers' retweeting behavior. Therefore, data analysis was mainly threefold, including geostatistical analysis and visualization, negative binomial regression analysis, and the mean comparison analyses through t-test and ANOVA.

3.1. Data collection

Data was collected from Twitter (www. twitter.com) from July 16th to 22nd, 2018, through the Twitter application programming interface (API). The selected tweets were scraped which included the key words. Restaurant Business (2018) names the top 50 profitable restaurant brands in the U.S. (e.g. McDonald's, Starbucks, Subway), which were used as keywords. A total of 22,207,502 tweets were collected. However, it should be noted that a tweet which includes a restaurant name may not focus on the content related to the restaurant experience. Several examples are listed below.

"I am a singer not a Starbucks drinker"

"how burger king still active lmao"

"There is a cop onjust before whataburger under the bridge"

Therefore, the initial tweet dataset was further screened with 29 keywords related to 29 restaurant and dining experience (e.g., dine-in, waiter, menu) developed by three experts in the hospitality and restaurant field. Through the preliminary and secondary screening of tweets, a total of 866,923 pieces of tweets written in English were considered in the data analysis process.

المنارة للاستشارات

www.manaraa.com

3.2. Text cleaning

Although using keywords largely increased the accuracy of identifying the specific customer tweets within the dinning context, a variety of redundant information was still kept in the customer tweets such as http link, referring to someone's information (@someone) and location. Since one of the primary purposes in the current study was to analyze textual content (emotions extracted from texts) of tweets posted by customers, it was necessary for taking text cleaning process before conducting multi-dimensional emotional analysis. Advised by Miner et al. (2012), the text cleaning process in the current study was followed with several steps, including removing punctuations and special characters, clearing numbers, escaping HTML links, and many others. And Table 7 demonstrates data cleaning criteria with sample words.

Table 7. Removed information and stop words in data cleaning

Removed information	Sample
Punctuation	",", ".", "?", ":"
Special character	"/", "@", "\\", "|"
Number	"1", "2", "3", "4"
Link	"http", "www", "com"

3.3. Text analysis

3.3.1. Linguistic Inquiry and Word Count

Customer review sentiments, emotions, and LSM were all detected and analyzed by using the statistic software Linguistic Inquiry and Word Count (LIWC) 2015. Introduced by Pennebaker et al. (2015), LIWC was originally developed to analyze individuals' verbal and written speech samples in an efficient manner. By calculating the percentage of a certain word and comparing with its own default dictionary, individual

www.manaraa.com

score value of each category such as cognitive, social, and function word categories was gained and calculated.

3.3.2. Multi-dimensional emotional framework

Referring to Pennebaker et al. (2015), although the default LIWC 2015 dictionary somehow involved inspections from both the perspectives of sentiments (i.e., positive and negative) and emotions (i.e., anxiety, anger, and sadness), an integrated multi-dimensional emotional framework as discussed in Section 2.3 cannot be directly recognized. Therefore, in order to detect eight emotions in tweets, the Word-Emotion Association Lexicon (EmoLex) dictionary developed by National Research Council (NRC) Canada was selected to satisfy the research purpose.

Introduced by Mohammad and Turney (2013), EmoLex can perfectly cover all of the eight emotional dimensions (i.e., anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) as well as two sentiments (i.e., negative and positive) in Plutchik's emotion wheel. Sample words under emotional dimension are provided in Table 8. With careful concerns of the most frequent English nouns, verbs, adjectives, and adverbs, EmoLex increased the quality of measuring the eight emotions from text materials. And in addition to a unigram dictionary provided by LIWC, EmoLex overcame the challenge of precisely annotating emotions not only from a single word (unigram), but also from a more sophisticated phrase or sentence (bigram) (Mohammad &Turney, 2010). Therefore, the EmoLex dictionary was imported into LIWC and applied to the multi-dimensional emotion analysis based on Plutchik' emotion wheel.

المنارة للاستشارات

www.manaraa.com

Table 8. Sample words of emotional dimensions in EmoLex

Emotional dimensions	Sample words
Anger	abuse, blame, confine, deny, evil, fraud
Anticipation	Adore, bonus, continue, develop, engaged, fortune
Disgust	Ashamed, badness, cheat, dirty, enemy, failure
Fear	Aggressive, bearish, captive, depress, endanger, fright
Joy	Accomplish, beauty, cheer, delight, esteem, fancy
Sadness	Absent, barren, cancel, damage, embarrass, flaw
Surprise	Amaze, bizarre, camouflage, dynamic, erupt, freakish
Trust	Admire, believe, confide, depend, electorate, faithful

3.3.3. Language Style Matching

Besides eight kinds of emotional dimensions, a total of LSM scores calculated from nine components was another variable of interest in the current study. As discussed above in Section 2.4, the LSM score measures the overall verbal coordination from nine independent function word categories. Therefore, in order to calculate the overall LSM score, a series of pre-procedures should be conducted based on Ireland and Pennebaker (2010). First, LIWC was used to identify and compute the score value of each of the nine LSM components; second, the LSM score formula was used to calculate LSM for each component. Taking the functional word category “conjunctions” as an example, referring to the formula below, the $conj_i$ refers to the conjunction score of a certain tweet provided by LIWC, while the $conj_G$ refers to the average score of the whole conjunction group; Third, the over LSM score was calculated with taking the average of nine components’ scores. Both individual LSM component score and the overall LSM score should range from 0 to 1, in which the degree of language style matching would rise with an increase in the overall LSM score (Gonzales, Hancock, & Pennebaker, 2010).

$$LSM_{conj} = 1 - [(|conj_i - conj_G|) \div (conj_i + conj_G + 0.0001)]$$

المنارة للاستشارات

www.manaraa.com

3.5. Inverse distance weighting (IDW)

As discussed in Section 2.3, emotions were assumed with geographic relationships, and customers' emotions expressed in their tweets also can both influence and be influenced by the surrounding tweets. However, after plotting all of the tweets with geographic information on the U.S. map showing in Figure 2, it is clear to see that most of the tweets only concentrated in large cities such as Los Angles, New York City, and Chicago, while a few of them speared in the Mid-western areas. Therefore, in order to have an integrated investigation on emotions in tweets from the two perspectives of both long-distance and short distance ranges, the inverse distance weighting (IDW) method was conducted in the current study only focusing on customers who provided the geographic information along with their tweets.

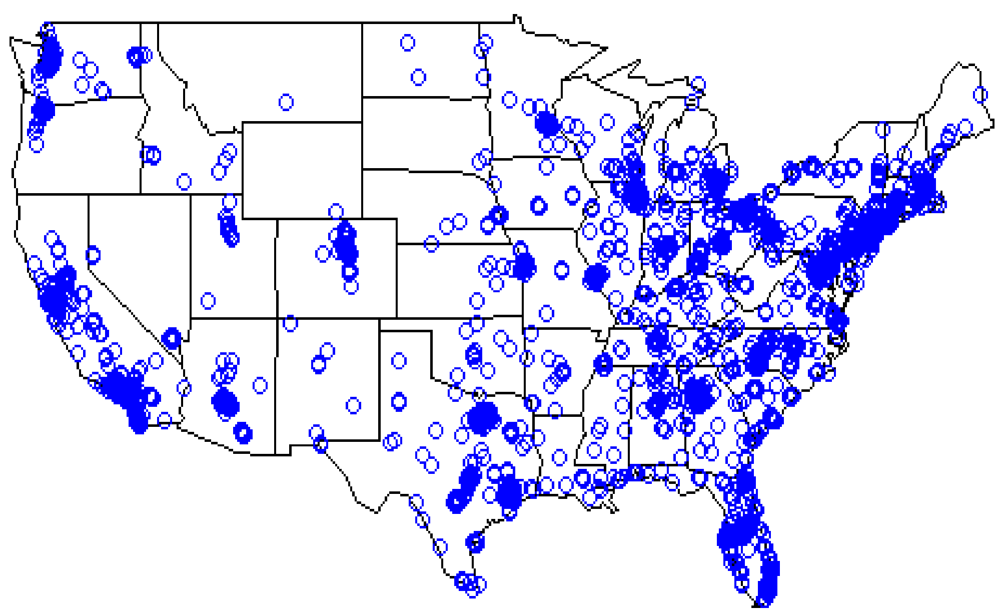

Figure 2. Geographic distribution of tweets

المنارة للاستشارات

www.manaraa.com

Within the geostatistical domain, the IDW is a kind of interpolation method to predict the unmeasured locations over an entire region by emphasizing an average of all nearby values (Azpurua & Ramos, 2010). With an assumption of Tobler (2004, p. 306)'s First Law of Geography that "everything is related to everything else. But near things are more related than distant things." IDW converts the geographic influence through mearing weight based on distance. Comparing with other traditional statistical methods such as taking average or using the variance, the IDW method can not only measure the geographic influence within a short distance, but also give consideration to the sample with a wide range by changing the power in the algorithm formula. And determined by the distance, a higher weight was calculated with those nearby observations, while a lower weight was calculated with those observations far away from a certain prediction point (Azpurua & Ramos, 2010).

3.6. Regression model

Identifying the determinants on customer retweeting behavior especially from the emotional aspect is the third purpose of the current study. And in order to focus on the retweeting behavior, customers' original tweets with at least one retweet (≥1) were kept in the new dataset. Since the range of the dependent variable was very huge referring to boxplots showing in Figure 3. In the current analysis, the value of 10,000 was manually set as the maximum for all count variables including the dependent variable of retweet number and the independent variables of user favorite count and user status count, so as to increase the precision of each measure and regression (Gardner et al., 1995), and finally the dataset contained 243,129 pieces of information for regression. The descriptive results of all variables of interest in the regression are presented in Table 9.

www.manaraa.com

Table 9. Descriptive analysis of variables

	Minimum	Maximum	Mean	Std. Deviation
Retweet number	1	9959	1259	2079
Positive	0.00	100.00	5.04	6.85
Negative	0.00	100.00	3.04	5.63
Anger	0.00	66.67	1.68	4.29
Anticipation	0.00	75.00	2.29	5.22
Disgust	0.00	50.00	0.93	3.11
Fear	0.00	50.00	1.42	3.69
Joy	0.00	80.00	2.62	5.26
Sadness	0.00	66.67	1.76	4.52
Surprise	0.00	50.00	1.02	3.30
Trust	0.00	100.00	2.49	4.99
LSM	0.00	0.82	0.30	0.18
Word count	0	33	13	6
Tweet time	0	23	12	8
Tweet weekday	1	7	4.14	2.06
User created year	2006	2017	2014	2
User favorite count	0	10000	3011	2639
User status count	0	10000	3360	2743
User verified	0	1	0.00	0.05

Note: n=243129

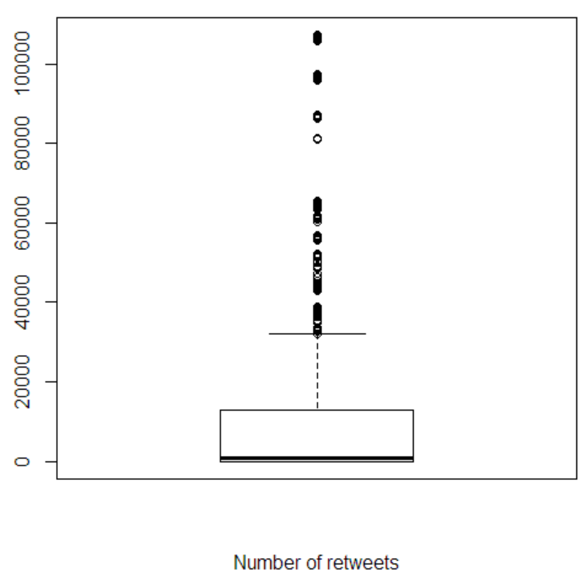

Figure 3. Boxplot of dependent variable

المنارة للاستشارات
www.manaraa.com

3.6.1. Dependent variable

The number of retweets (i.e., how many times the original tweet was reposted) was used as the dependent variable in the current analysis. According to the descriptive analysis in Table 36, the range of the dependent variable is 9,958 with a minimum value at 1, and maximum value at 9959. And the distribution of the dependent variable of retweet number is provided in Figure 4. Based on the theory of Elaboration Likelihood Model (ELM) discussed in Section 2.3, when customers read or review some fragmental pieces of information posted on online social media, they would follow the peripheral route of persuasion, where emotions expressed along with their textual contents can also influence customers' decision-making process (Henningsen et al., 2003). Therefore, the retweet number was used to represent customer preference on a certain tweet and also reflect their retweeting intention.

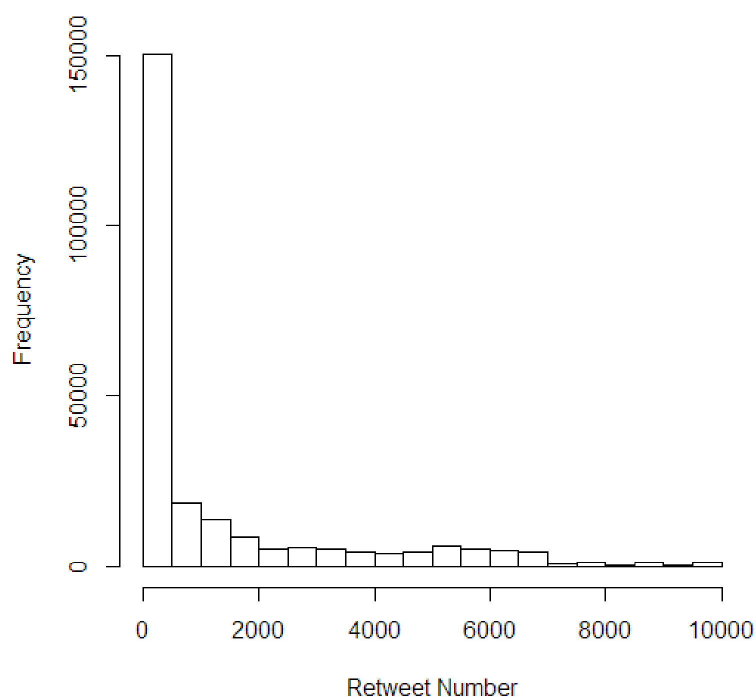

Figure 4. Distribution of retweet number

المنارة للاستشارات
www.manaraa.com

3.6.2. Independent variable

Independent variables of interest considered in the current analysis included two sentiments, eight emotional dimensions, and the overall LSM score calculated from customer tweets. According to the Plutchik (1994)'s emotion wheel, two sentiments were consisted of positive and negative emotion polar, while the eight emotional dimensions were consisted of four opposite pairs of emotions, including sadness-joy, fear-anger, disgust-trust, and surprise-anticipation. All value scores of both sentiments and emotions were calculated by LIWC with a range from 1 to 100. The distribution of two sentiments is provided in Figure 5, the distribution of four positive emotional dimensions is demonstrated in Figure 6, and the distribution of four negative emotional dimensions is shown in Figure 7.

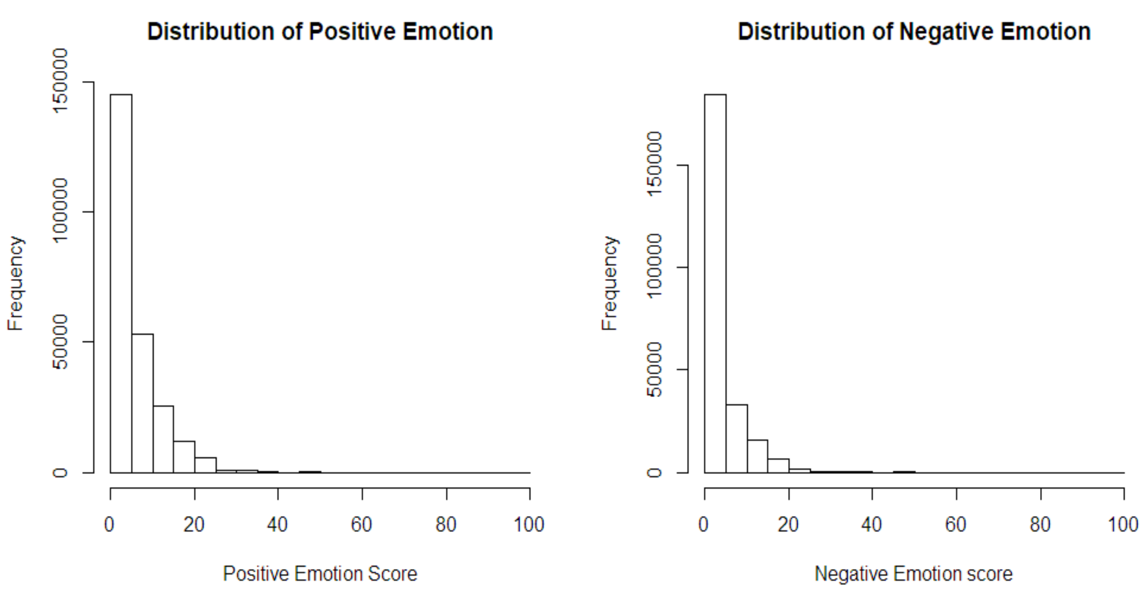

Figure 5. Distribution of two sentiments

المنارة للاستشارات

www.manaraa.com

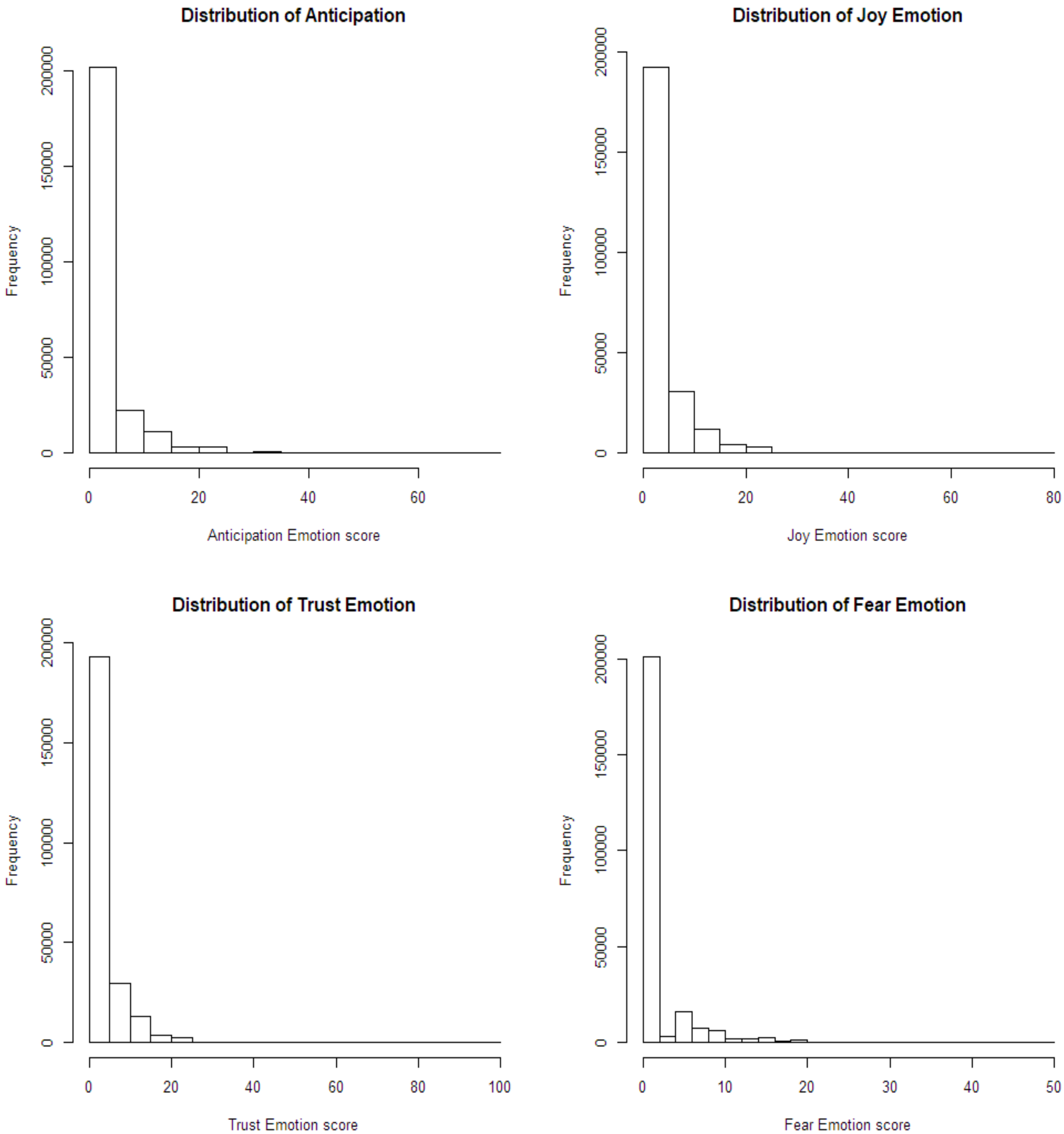

Figure 6. Distribution of four positive emotions

www.manaraa.com

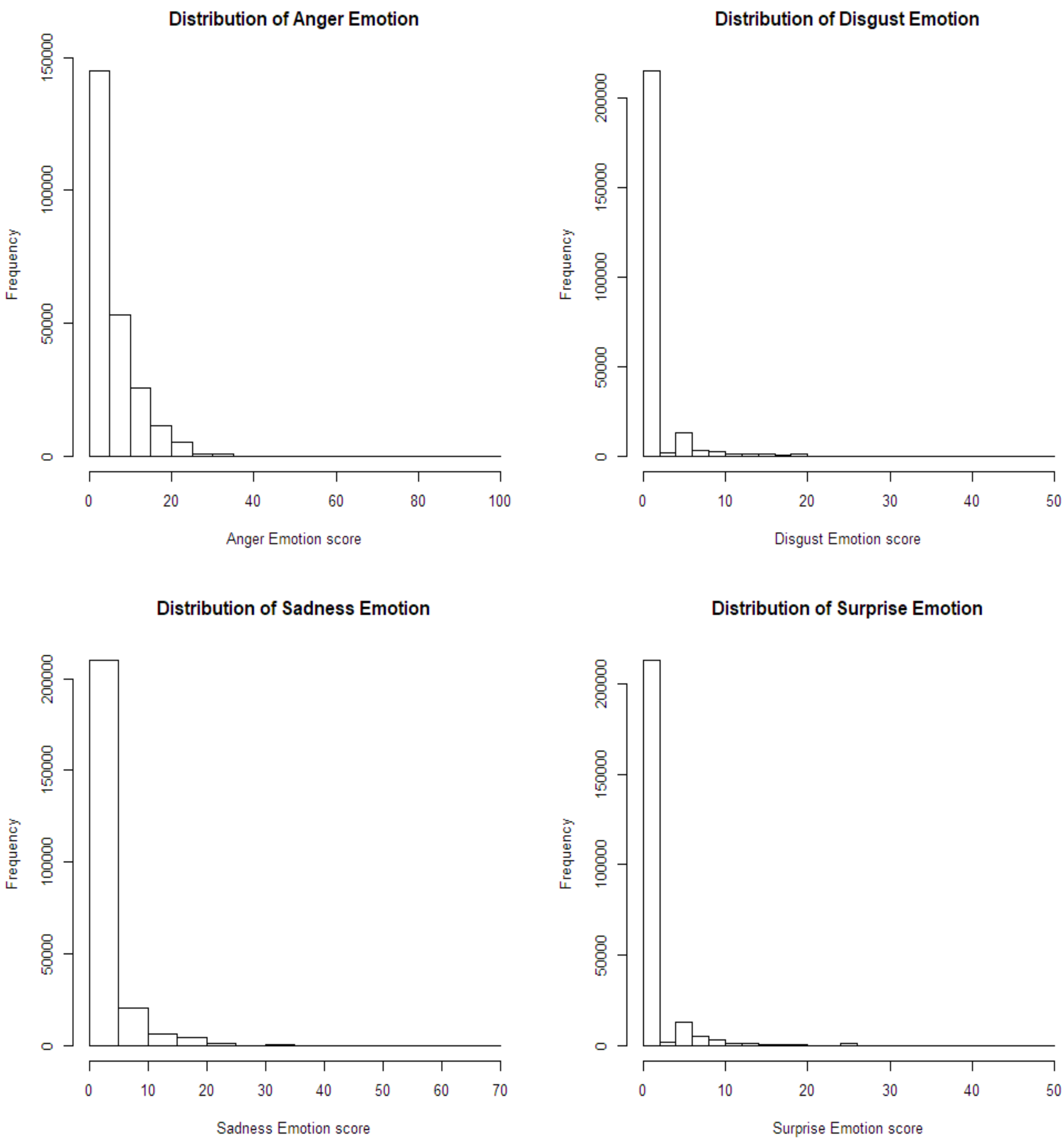

Figure 7. Distribution of four negative emotions

www.manaraa.com

And as mentioned in Section 3.3, LSM was used to measure the matching of function words, and the overall LSM score was calculated by taking the average value of all nine sub-components of the functional words. With a range from 0 to 1, the distribution of LSM is demonstrated in Figure 8.

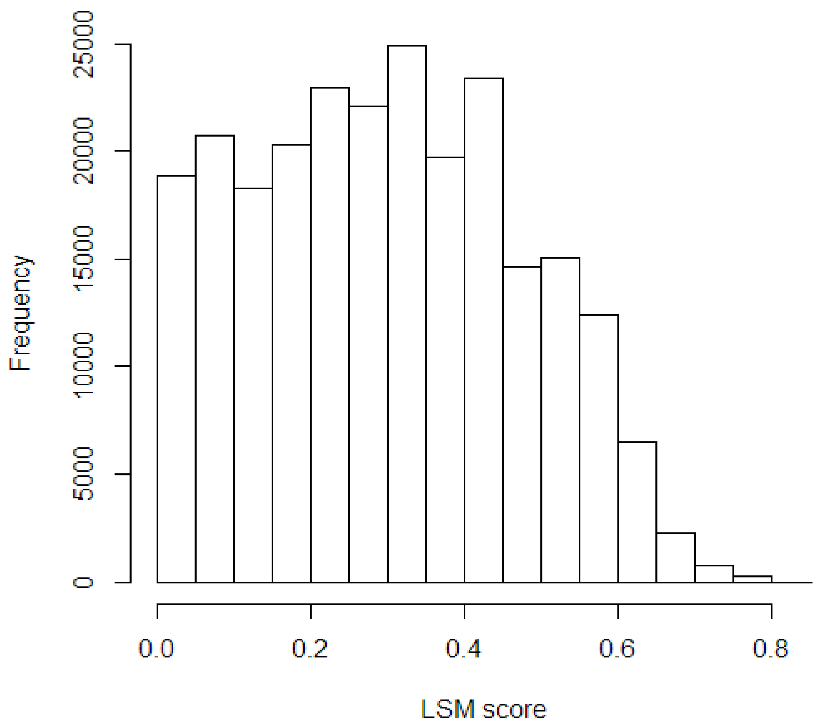

Figure 8. Distribution of the overall LSM score

3.6.3. Control variable

Control variables contained in the current analysis were mainly concerned about three attributes related to the tweets and three customer profile variables associated with the account. As to the attributes related to tweets, word count indicated word number in a certain tweet; tweet time showed what time was the tweet created that was measured in a 24-hour time frame; and tweet weekday indicated which day of the week from one to

www.manaraa.com

seven the tweet was created. As to the customer profile, user created year indicated which year the user account was created with a range from 2006 to 2017; user favorite count showed how many pieces of information the account was favorited; and user status count represented how many tweets the account was posted. In addition, one more dummy variable, user verified was considered in the current study as well. Measured by “0″ and “1″ referring to the Yes and No, respectively, the dummy variable indicated if a certain user is officially verified by the Twitter company.

3.6.4. Negative binomial model

Negative binominal model was selected to conduct the regression analyses in the current study mainly for two purposes. On the one hand, according to the main distributions of both dependent and independent variables presented above in Figure 4-8, it was clear to identify that not only the data of dependent variable retweet number appeared as right-skewed distribution rather than normally distributed, but also the data of independent variables including two sentiments and eight emotional dimensions were also right-skewed distributed. Therefore, the negative binomial regression was a proper method to model these right-skewed distributions of dataset (Fang et al., 2016). On the other hand, in terms of the attribute of the dependent variable of retweet number, it met the standard of count model in which the retweet number was actually counted from 1 to the positive infinity (Gardner et al., 1995).

Therefore, as an extended model from the Poisson regression, the negative binomial regression was designed to model the count distribution of dataset as well, and at the same time, it can be applied to process the analysis more precisely than the Poisson regression on the over-dispersed count data that the mean and variance are not the same (Table 9). Moreover, superior to traditional ordinary least squares (OLS) regressions, the

www.manaraa.com

negative binomial regression was regarded as one of the particular methods in measuring dependent variables which were counted as nonnegative integers (Gardner et al., 1995). Negative binomial model was applied in the current analysis.

www.manaraa.com

CHAPTER 4. RESULTS

4.1. Correlation

Pearson correlation results of all variables are presented in Table 10. According to the results, the dependent variable of retweet number was identified with the relationship among all independent variables and control variables, which provided supports to conduct further regression analyses in the next section. On the other hand, based on the correlation table, there was still some multi-collinearity relationship among independent variables that can be identified. Therefore, the variance inflation factor (VIF) was conducted as a supplementary test to quantify the severity of multicollinearity in the next section as well (O'brien, 2007).

4.2. Negative binomial regression

One of the primary purposes of the current study was to investigate the effect of both sentiment and emotions on customer retweeting behavior. Therefore, rather than putting all variables into one regression, the current study separately measured the two polar of sentiments (i.e., positive and negative) in model 1, while measured the eight emotional dimensions (i.e., anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) in model 2. In the meantime, language style matching (LSM score) was also taken into consideration in both two models. And the results of two negative binomial regression models are presented in Table 11 and Table 12.

المنارة للاستشارات
www.manaraa.com

Table 10. Correlation matrix of variables

	1	2	3	4	5	6	7	8	9
1 Retweet number	1								
2 Positive	0.055**	1							
3 Negative	-0.024**	-0.072**	1						
4 Anger	-0.034**	-0.039**	0.743**	1					
5 Anticipation	0.089**	0.426**	0.014**	0.019**	1				
6 Disgust	-0.044**	-0.010**	0.504**	0.476**	0.010**	1			
7 Fear	0.027**	0.017**	0.491**	0.469**	0.035**	0.502**	1		
8 Joy	0.059**	0.685**	-0.053**	-0.027**	0.506**	0.009**	0.017**	1	
9 Sadness	-0.073**	0.027**	0.676**	0.662**	0.025**	0.445**	0.448**	-0.014**	1
10 Surprise	0.026**	0.300**	0.051**	0.081**	0.409**	0.077**	0.099**	0.443**	0.079**
11 Trust	0.097**	0.546**	-0.003	0.006**	0.470**	0.004	0.044**	0.553**	0.017**
12 LSM	0.080**	-0.051**	0.026**	-0.013**	-0.049**	0.053**	0.051**	-0.044**	-0.049**
13 Word count	0.087**	-0.00337	0.024**	-0.020**	-0.024**	0.045**	0.060**	-0.023**	-0.055**
14 Tweet time	-0.008**	0.015**	-0.011**	-0.009**	0.012**	-0.004*	0.004	-0.001	-0.010**
15 Tweet weekday	-0.025**	-0.011**	0.019**	0.022**	-0.015**	0.022**	0.005**	0.015**	0.025**
16 User created year	-0.025**	-0.012**	-0.011**	0.007**	-0.004	-0.001	-0.011**	-0.004*	0.003
17 User favorite count	0.038**	-0.001	0.013**	0.010**	0.001	0.001	0.000	0.017**	0.005**
18 User status count	0.024**	0.006**	0.017**	0.016**	-0.002	0.007**	0.007**	0.002	0.011**
19 User verified	-0.028**	0.006**	-0.003	-0.003	0.005*	-0.004*	0.002	-0.001	-0.007**

Note: Two-tailed Pearson correlation. **p<0.01; *p<0.05

المنارة للاستشارات
www.manaraa.com

Table 10. Correlation matrix of variables (continued)

	10	11	12	13	14	15	16	17	18	19
10 Surprise	1									
11 Trust	0.364**	1								
12 LSM	-0.033**	-0.007**	1							
13 Word count	0.008**	0.038**	0.814**	1						
14 Tweet time	0.005*	0.007**	0.010**	0.026**	1					
15 Tweet weekday	-0.007**	0.001	-0.022**	-0.018**	0.001	1				
16 User created year	0.006**	-0.035**	-0.026**	-0.033**	-0.016**	0.019**	1			
17 User favorite count	0.006**	-0.007**	0.021**	0.002	-0.004*	-0.002	-0.136**	1		
18 User status count	-0.004	0.001	-0.014**	-0.021**	-0.002	0.004*	-0.295**	0.436**	1	
19 User verified	-0.001	0.003	0.002	0.009**	0.014**	-0.006**	-0.081**	-0.011**	0.027**	1

Note: Two-tailed Pearson correlation. ***p<0.01; **p<0.05

للاستشارات المنارة
www.manaraa.com

In model 1, with a 2 × log-likelihood value at -3290473.0800 and AIC value at 3290497, all independent and control variables were significant. According to the correlation table, some correlations were shown between independent variables. Therefore, the variance inflation factor (VIF) test was provided. VIF in model 1 was ranged between 1.0015 and 2.9837, which indicated that the multicollinearity effect between variables was acceptable with a standard cut-off value less than 5 (Hair et al., 1998).

As to the independent variables, positive sentiment ($\beta=0.0136$, $p<0.0000$) showed a positive effect on the customer retweet number which indicated positive sentiments expressed in customers' tweets play an vita role to motivate retweet behavior and increase retweet number, while negative sentiment ($\beta=-0.0134$, $p<0.0000$) showed a negative effect on retweet number, which indicated negative sentiment expressed in customers' tweets can diminish the retweeting desire of customers and reduce the retweet number. In the meantime, LSM ($\beta=0.2172$, $p<0.0000$) showed a positive effect on the customer retweet number, which indicated higher stylistic similarity between two users can motivate retweeting behavior.

As to the control variables, word count ($\beta=0.0189$, $p<0.0000$), user favorite count ($\beta=0.000$, $p<0.0000$), and user status count ($\beta=0.0000$, $p=0.0174$) had positive effects on customer retweet number; but tweet time ($\beta=-0.0030$, $p<0.0000$), tweet weekday ($\beta=-0.0077$, $p=0.0004$), user created year ($\beta=-0.0132$, $p<0.0000$), and user verified ($\beta=-2.2456$, $p<0.0000$) had negative effects on retweet number.

المنارة للاستشارات
www.manaraa.com

Table 11. Negative binomial regression results focusing on sentiments

Variables	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	33.2985	4.2904	7.7610	0.0000	***
Positive	0.0136	0.0007	20.8200	0.0000	***
Negative	-0.0134	0.0008	-16.8850	0.0000	***
LSM	0.2172	0.0439	4.9440	0.0000	***
Word count	0.0189	0.0012	15.9260	0.0000	***
Tweet time	-0.0030	0.0006	-5.1920	0.0000	***
Tweet weekday	-0.0077	0.0022	-3.5540	0.0004	***
User created year	-0.0132	0.0021	-6.1870	0.0000	***
User favorite count	0.0000	0.0000	11.4010	0.0000	***
User status count	0.0000	0.0000	2.3790	0.0174	**
User verified	-2.2456	0.0854	-26.2910	0.0000	***

Note: ***p < 0.001; *p< 0.05

As to model 2, with a 2 × log-likelihood value at -3286807.1780 and VIF value ranged between 1.0018 and 2.9980, all independent variables were significant, and all control variables excluding tweet weekday were significant. Comparing with the two models, although log-likelihood values cannot be used alone to directly measure the model fit, and the higher value is better. Therefore, when applying in-depth analysis reaching the individual emotional dimension, with a 3,665.9020 increased in the 2 × log-likelihood value, model fit of the second regression increased, which also contributed to the purpose of the current study which used the multi-dimensional emotional framework to do the analysis. In addition, the comping with model 1, AIC value of model 2 also decreased to 3286843 and proved with a better model fit in model 2 as well.

As to the independent variables, anticipation (β=0.0233, p<0.0000), fear (β=0.0752, p<0.0000), and trust (β=0.0189, p<0.0000) showed significant positive effects on customer retweet number, which revealed emotional dimensions of anticipation, fear, and trust can positively motivate customer retweeting behavior and increase the retweet number; however, anger (β=-0.0059, p=0.0001), disgust (β=-0.0470, p<0.0000), joy (β=-0.0062, p<0.0000),

المنارة للاستشارات

www.manaraa.com

sadness (β=-0.0607, p<0.0000), and surprise (β=-0.0044, p=0.0053) showed negative effects on customer retweet number, which revealed emotional dimensions of anger, disgust, joy, sadness, and surprise can reversely influenced customer retweeting behavior and decreased the retweet number. LSM (β=0.3604, p<0.0000) also showed a positive effect on the customer retweet number, and also proved that a higher stylistic similarity between two users can motivate retweeting behavior.

As to the control variables, word count (β=0.0113, p<0.0000), user favorite count (β=0.000, p<0.0000), and user status count (β=0.0000, p=0.0177) had positive effects on customer retweet number; but tweet time (β=-0.0043, p=0.0006), user created year (β=-0.0108, p<0.0000), and user verified (β=-2.3092, p<0.0000) had negative effects on retweet number. Tweet weekday did not show a significant effect in this model.

Table 12. Negative binomial regression results focusing on emotions

Variables	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	28.5155	4.2704	6.6780	0.0000	***
Anger	-0.0059	0.0015	-4.0340	0.0001	***
Anticipation	0.0233	0.0010	22.2530	0.0000	***
Disgust	-0.0470	0.0018	-26.7270	0.0000	***
Fear	0.0752	0.0015	50.7290	0.0000	***
Joy	-0.0062	0.0011	-5.5630	0.0000	***
Sadness	-0.0607	0.0014	-44.7010	0.0000	***
Surprise	-0.0044	0.0016	-2.7860	0.0053	**
Trust	0.0189	0.0011	16.9270	0.0000	***
LSM	0.3604	0.0437	8.2460	0.0000	***
Word count	0.0113	0.0012	9.5710	0.0000	***
Tweet time	-0.0043	0.0006	-7.4060	0.0000	***
Tweet weekday	0.0032	0.0022	1.4960	0.1348	
User created year	-0.0108	0.0021	-5.0980	0.0000	***
User favorite count	0.0000	0.0000	11.9810	0.0000	***
User status count	0.0000	0.0000	2.3710	0.0177	*
User verified	-2.3092	0.0849	-27.1830	0.0000	***

Note: ***p < 0.001; **p<0.01; *p< 0.05

المنارة للاستشارات

www.manaraa.com

4.3. Emotion timeline analysis

Besides only concerning the independent variables, one of the control variables – tweet time, captured the attention in the current study as well. Identified with a significant effect on customer retweeting behavior in both two regression models, in this section, I further employed an investigation of sentiments and emotions expressed along with tweets based on a timeline sequence.

In the current analysis, the total timeline was broken down into six time periods with a four-hour interval. Rather than simply beginning with 12:00am, referring to the people dining habit, 2:00am was set as the beginning value, which can be more accurate to specify each customer dinning period including the early morning (2:00am-5:00am), breakfast time (6:00am-9:00am), lunchtime (10am-1:00pm), afternoon (2:00pm-5:00pm), dinner time (6:00pm-9:00pm), nightlife (10:00pm-1:00am).

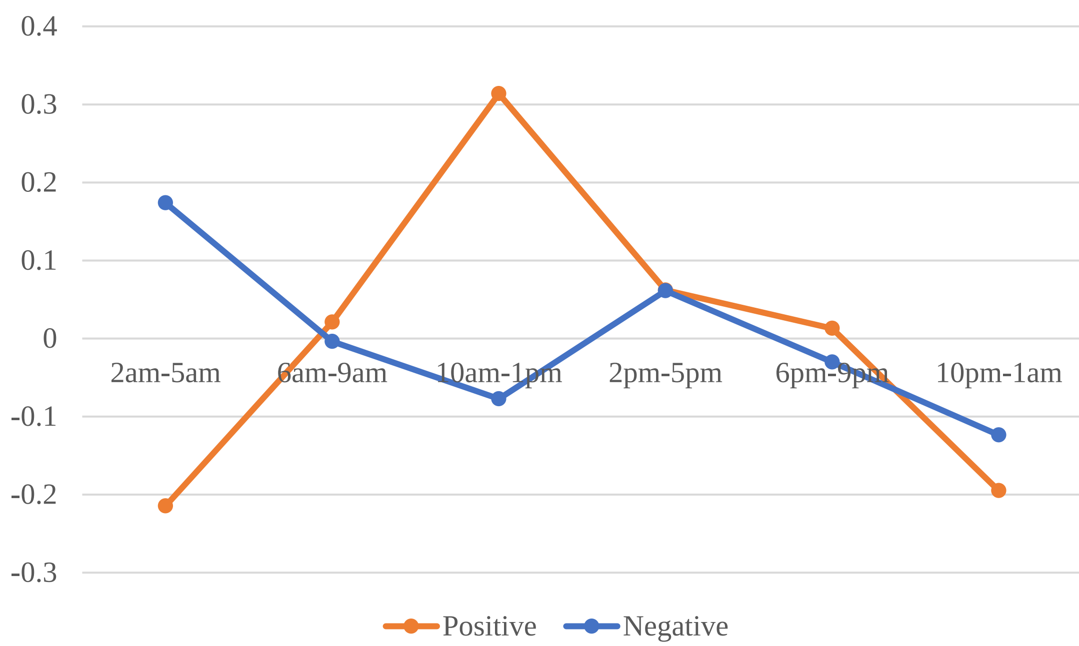

Figure 9. Timeline analysis of two sentiments

المنارة للاستشارات
www.manaraa.com

The timeline analysis of two sentiments is presented in Figure 9. After a standardized process, the sentiment expressed in a certain time period, the sentiment trend along with a timeline, and the comparison between two sentiments can be clearly identified. As to the positive sentiment group, positive sentiment expressed in tweets increased from 2:00am to 1:00pm and reached to the peak at 10:00am - 1:00pm period. After 1:00pm, the positive sentiment value dropped down. However, the negative sentiment reached to the peak at the beginning at 2:00am - 5:00am time period, and then, the negative emotions expressed around the day were relatively bland.

Through the comparison between positive and negative sentiment group, indicated from Figure 9, positive emotion hit to bottom, but negative emotion reached to the peak at 2:00am-5:00am, while positive emotion reached to the peak, but negative emotion decreased near the bottom at 10:00am - 1:00pm. Within the 2:00am to 1:00pm period, positive emotion increased but negative emotion decreased. Within the 2:00pm to 1:00am period, both positive and negative emotions decreased correspondingly. However, when focusing on the beginning and the end of the timeline, it should be noticed a compelling difference that although negative emotion dropped to the bottom value at the 10:00pm - 1:00am time period, it extremely reached to the peak at the 2:00am - 5:00am period as mention before. And the results were also statistically supported from the t-test presenting in Table 13, which indicated a significant difference ($p<0.0000$) between two sentiment groups (i.e., positive and negative).

Table 13. Results of t-test of two sentiment groups

Group	N	Mean	Variance	df	t Stat	P(T<=t)	t-critical
Positive	6	5.1292	0.0374	8	26.0806	0.0000	2.3060
Negative	6	2.7815	0.0112				

www.manaraa.com

Besides, the current analysis also further explored the emotions expressed in tweets narrowing down to the eight dimensions. Based on the ANOVA results (Table 14), it also can be clearly identified significantly difference (p<0.0000) between eight emotions based on the timeline. In the meantime, a radar chart is provided in Figure 10 which accurately presented eight emotions expressed based on the timeline.

Table 14. ANOVA test of eight emotional dimensions

Group	SS	df	MS	F	P-value	F critical
Between group	17.5454	7	2.5065	390.8677	0.0000	2.2490
Within group	0.2565	40	0.0064			
Total	17.8019	47				

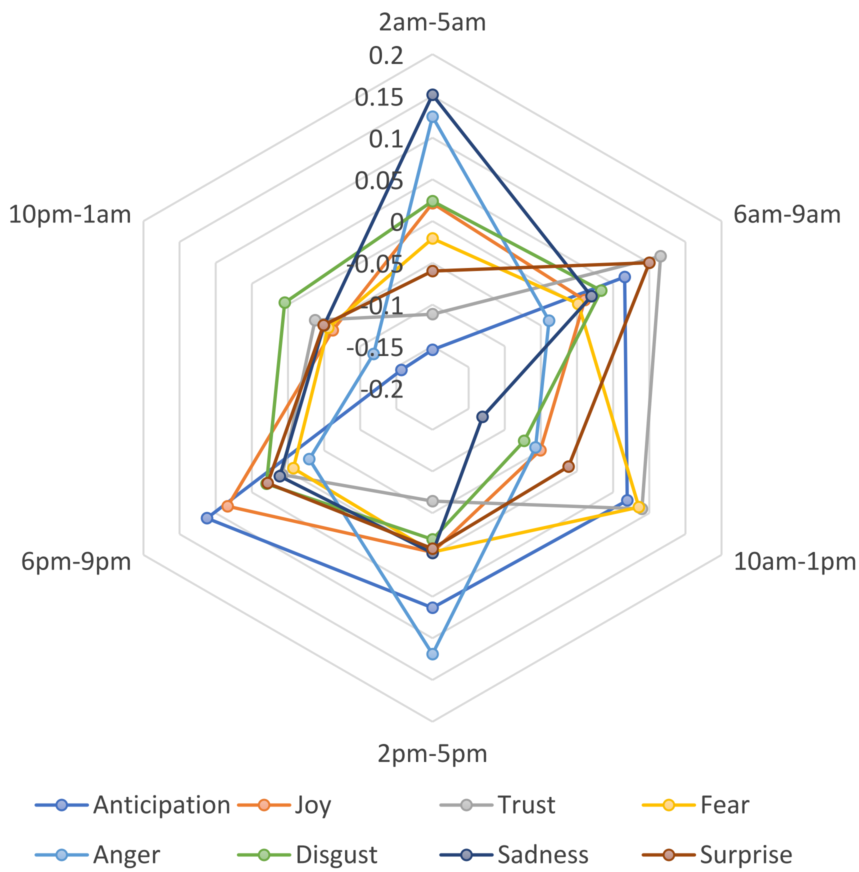

Figure 10. Eight emotions expressed based on timeline

المنارة للاستشارات
www.manaraa.com

4.4. Geostatistical analysis

Based on the negative binomial regression results, independent variables of interests of both two sentiments and eight emotional dimensions were identified having a significant influence on customer retweeting behavior. Therefore, in this section, the geostatistical analysis method was applied to have a statistical estimation based on the existing dataset and predict the spread of sentiments on those unknown areas.

Data used in the geostatistical analysis were extracted from the total customer tweets dataset. Besides the rigid requirement that the tweet should provide geographic coordinates of both longitude and latitude, several standards were also necessary when filtering those tweets, for example, the location of the posted tweet should be located within the U.S. boundary, and the tweets should also be retweeted by other customers. Therefore, under those strict conditions, there were 4,321 pieces of tweets gathered in the geostatistical analysis.

In the current study, geostatistical analysis only explored to the sentiment level formed the positive emotion polar and negative emotion polar, and those two sentiment levels of customer tweets are plotted in Figure 11. According to the plot figure, most of the tweets either in positive or negative sentiments were indicated at a slightly weak level with the score range from 0 to 10; although lower than the first level, there were still many sentiments with a score ranged from 10 to 20; but comparing with the first two levels, the number of tweets with a sentiment score larger than 20 is not many. And the frequency table of sentiments existing in the customer tweets is summarized in Table 15.

Table 15. Frequency of each sentiment level

Sentiment	[0,10]	(10,20]	(20,30]	(30,40]	(40,50]	Total
Positive	3,558	670	85	8	0	4321
Negative	4,162	141	17	0	1	4321

المنارة للاستشارات

www.manaraa.com

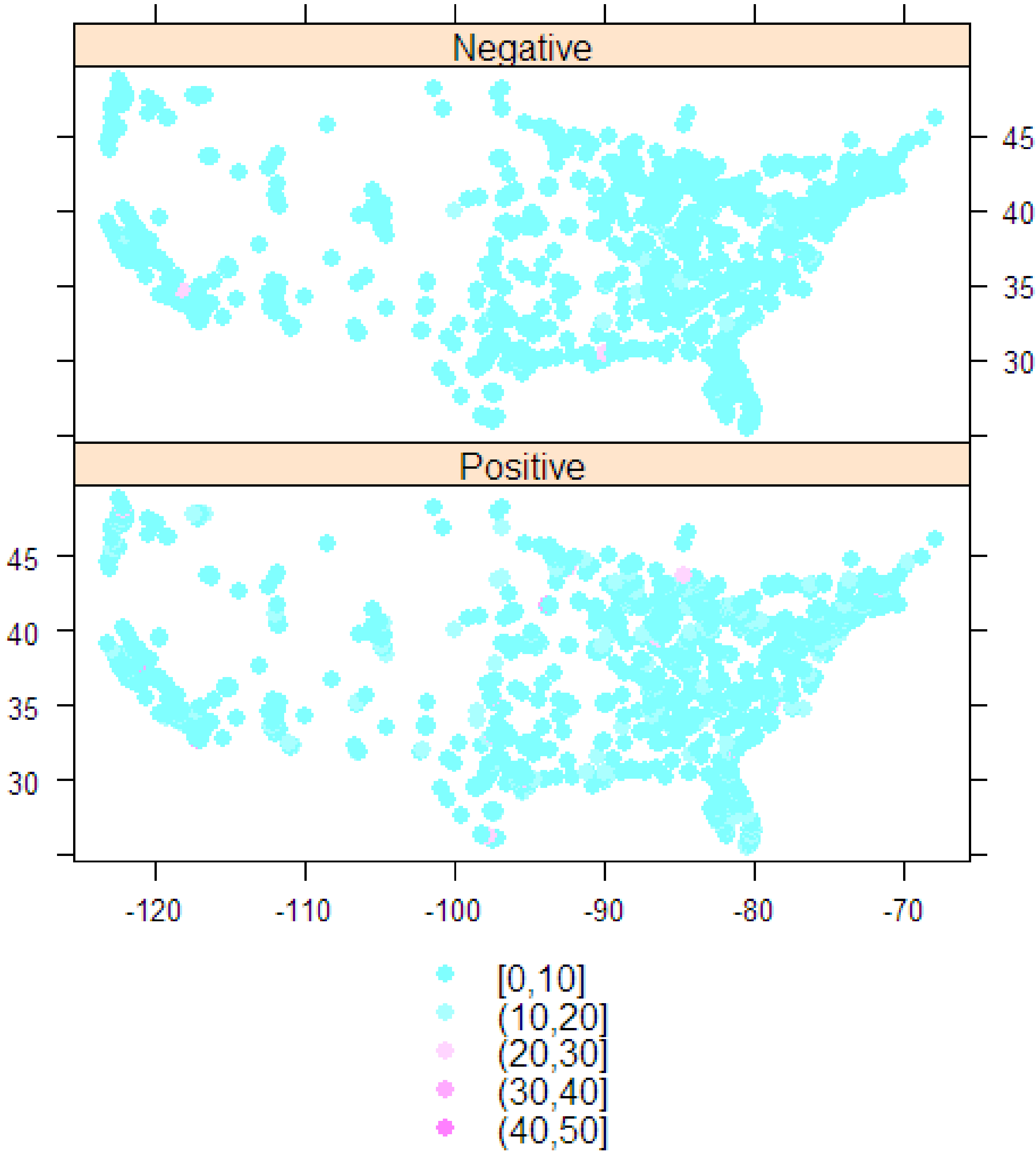

Figure 11. Sentiment level of customer tweets

www.manaraa.com

With a primary purpose to identify the sentiments in those unknown areas, inverse distance weighted interpolation method was applied in the estimating process. In the current analysis, the total areas of interest of the overall US boundary was separated into 80,000 grid areas including 400 grids in length and 200 grids in width, the total areas of interest of the overall California state was separated into 10,000 grid areas including 100 grids in length and 100 grids in width, and the total areas of interest of Los Angles city was separated into 2,500 grid areas including 50 grids in length and 50 grids in width.

The IDW analysis was conducted in R 3.6 with a "gstat" package, in which the power value was set at 1 when estimating sentiment scores in the US and sentiment scores in the California state, but the power value was set at 2 when estimating the Los Angeles city. Because based on the power function showing in Figure 12, the smaller the power is more appropriate to measure a longer distance, but the larger power value can be more appropriate to measure "more local" area with a short distance.

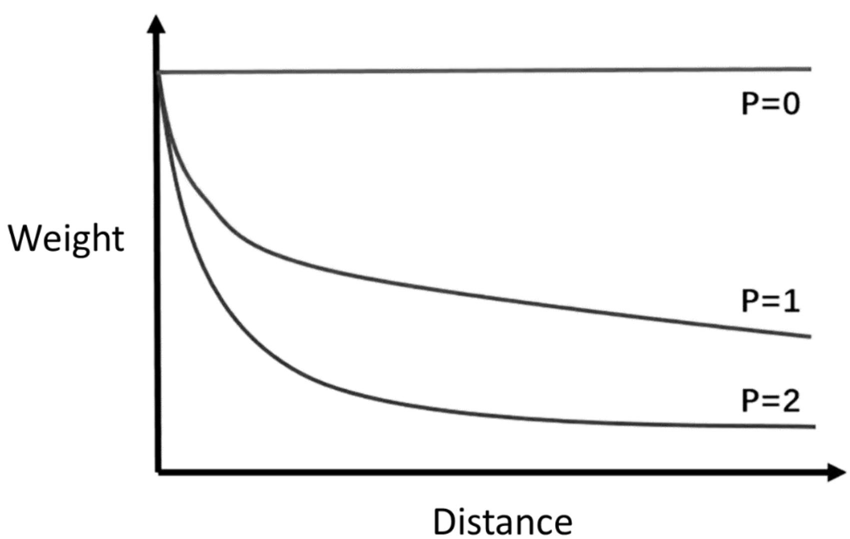

Figure 12. Power function of IDW

المنارة للاستشارات
www.manaraa.com

Geostatistical estimation results of both positive and negative sentiments are presents in Figure 13 and Figure 14. According to those figures, positive emotion in customers' tweets was primarily expressed in several areas around Seattle, Phoenix, St. Louis, New York City, and Miami, while the negative emotions were mainly expressed around Atlanta and Washington DC. Although those two figures provided an overall preview of the sentiments' estimation by IDW interpolation method, it only presented in a very larger area with a long-distance range.

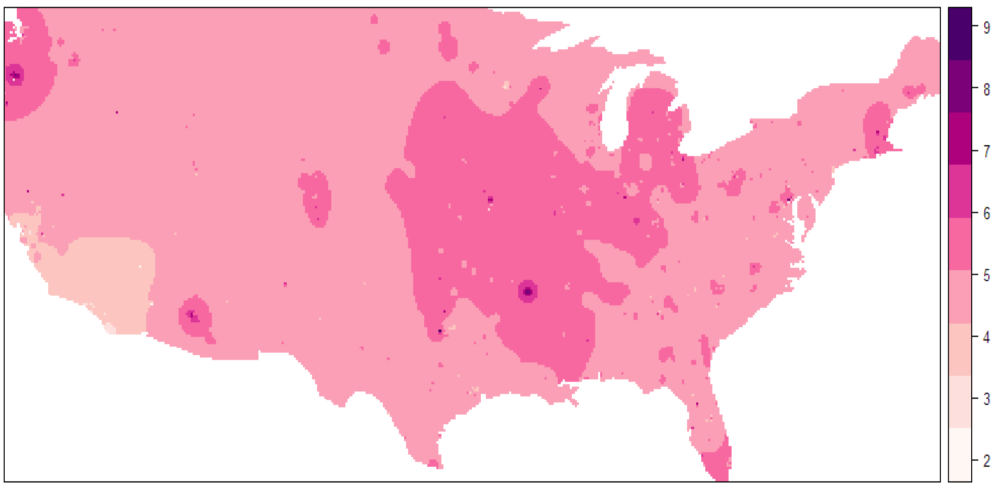

Figure 13. Positive sentiment estimation by IDW

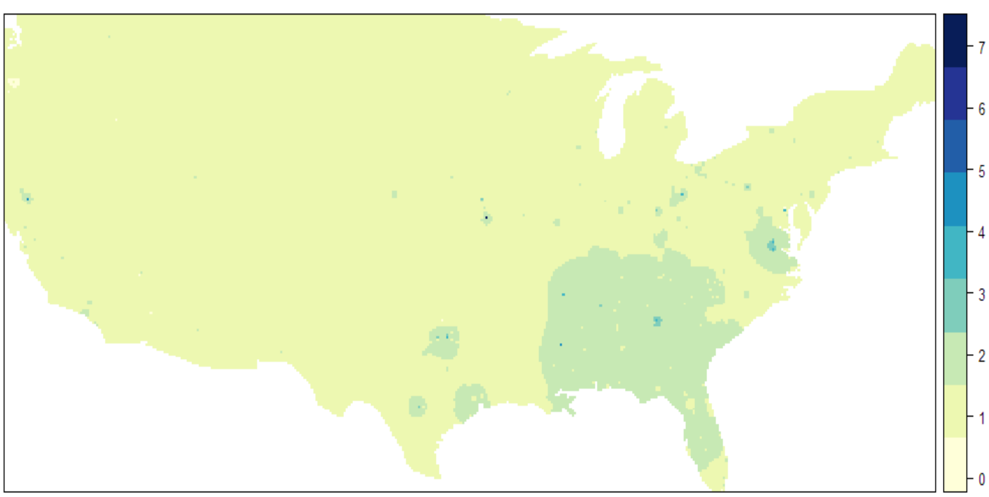

Figure 14. Negative sentiment estimation by IDW

www.manaraa.com

However, in order to have a more precise analysis of sentiments, it was necessary to further investigate the estimation within a shorter distance range such as a state or a city. Therefore, with this consideration, both California state and the Los Angeles city were chosen to the analysis in the next step. Because on the one hand, there are two major cities San Francisco and Los Angeles in California, it can clearly reflect the relationship between these two cities when doing IDW estimation. On the other hand, Los Angeles is one of the biggest cities in the US, the number of tweets there is sufficient to do the geostatistical analysis.

Results of IDW estimation in California state are presented in Figure 15 and Figure 16, and results in Los Angeles city are presented in Figure 17 and Figure 18. It was not hard to identify that positive sentiment spread wider than the negative sentiment, and positive sentiment spread with a higher level at around 3 to 5 than the negative sentiment at the California state level.

However, when referring to a narrower area of Los Angeles city, although customers were influenced by both positive and negative sentiments expressed in their tweets, estimation results were not spread in the same pattern. As to the positive sentiment estimation, positive sentiment was mainly expressed in the south part of Los Angeles, and it was indicated with a slight extension to other areas at a level of 3 to 4. But as to the negative sentiment estimation, north part of Los Angeles was the primary source expressing the negative sentiment, but the negative sentiment was spread almost round the entire city at a level between 1.5 to 2.

المنارة للاستشارات

www.manaraa.com

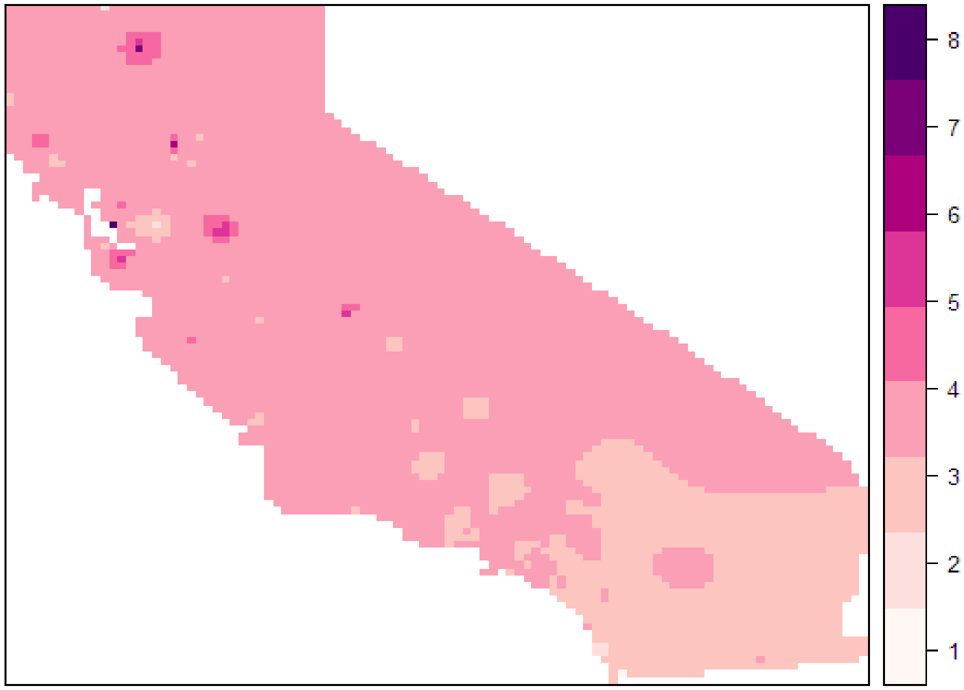

Figure 15. Positive sentiment estimation of California

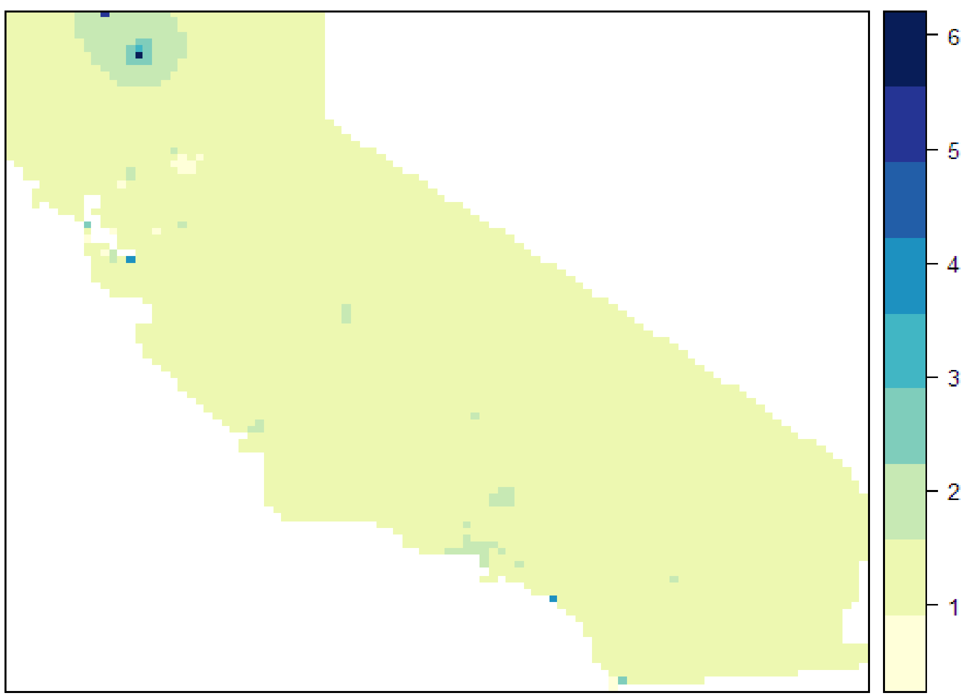

Figure 16. Negative sentiment estimation of California

www.manaraa.com

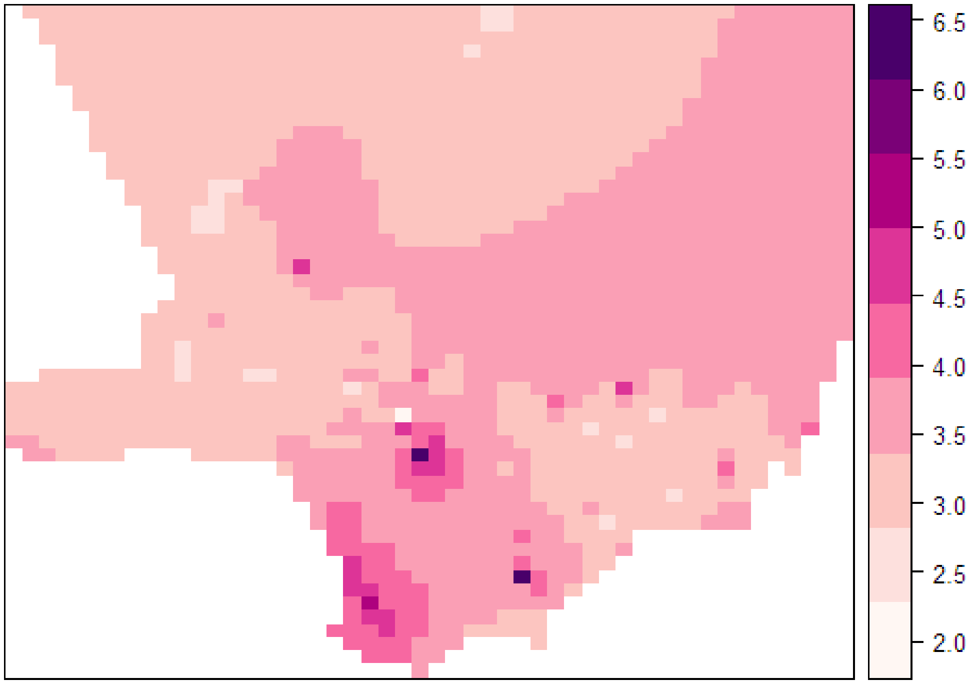

Figure 17. Positive sentiment estimation of Los Angeles

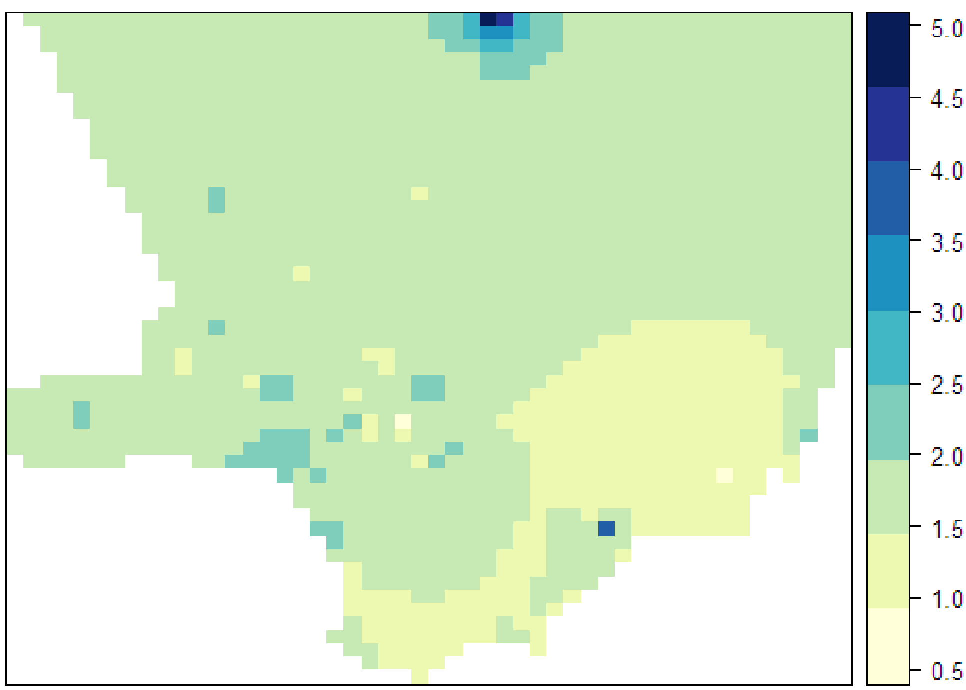

Figure 18. Negative sentiment estimation of Los Angeles

www.manaraa.com

CHAPTER 5. DISCUSSIONS

The current study investigated the effects of online UGC on customer retweeting behavior creatively through the text mining method. Through a series of calculation algorithms and computing procedures, variables of interests were transformed and measured by the score value. Data analysis results demonstrated all propositions in the current study, and proved that all of the independent variables have significant effects on the dependent variable of customer retweeting behavior.

Indicated in the negative binomial regression analysis, in model 1, positive sentiment along with the tweets posted by customers showed a positive influence on the customer retweeting behavior, while negative sentiment showed a negative influence. The analysis results related to both of the two sentiments were not surprising, because according to the discussion of Joyce and Kraut (2006), positive sentiments included in messages during communications can reinforce a sense of community and encourage continued participation, but negative sentiments normally result in hostile atmosphere and insult conversations between each other. Therefore, with a sense of positive sentiments of tweets, potential customers would increase the prospect to join in the conversations on the topic and share the information with other friends. However, tweets, as well as conversions expressed along with the negative sentiment, would more likely be stopped, and the communication session would be terminated without further diffusion to a larger group of people. And the results of the sentiment analysis consisted of previous literature like Duffett (2015) and Kwon and Choi (2014).

However, exploring the two-dimensional sentiments from positive and negative only can provide an ambiguous evaluation, which overlook the complex affection from many distinctive perspectives of emotions (Wang, Tang, & Kim, 2019). Therefore, the second regression model

المنارة للاستشارات
www.manaraa.com

conducted in the current made up the exhaustive analysis focusing on eight sub-emotions (four opposite pairs) under the two sentiments. According to the Plutchik (1994)'s emotional framework, joy, trust, fear, and anticipation are positive emotions, but surprise, sadness, disgust, and anger are negative emotions. In model 2, emotions including anticipation, fear, and trust indicate significant positive effects on customer retweeting behavior; while emotions including anger, disgust, sadness, surprise, and joy indicate significant negative effects on the dependent variable. Comparing with the result of sentiment analysis, the majority of emotions, except the aspect of joy, were identified with the same direction of influence with either of two sentiments.

As to the three emotions with positive effects, anticipation normally contained the expression of thoughts about the future, and it normally situated from interest to vigilance. Associating with the dinning situation, if tweets express with an anticipation emotion and show interests to a certain restaurant or the food and service, such a type of tweet information are advised to increase the retweet number. The emotion of fear normally situates from timidity to terror. Fear contained in tweets can express a strong feeling to the hygiene situation or dining environment, which can also grab people's attention and have a chain reaction such as bad-of-mouth among a group of people. Trust situated from acceptance to admiration, different from fear, emotion of trust can contribute to the electronic word-of-mouth so as to increase the popularity and amplify the reputation of a restaurant among people.

As to the five emotions with negative effects, anger situates from annoyance to fury, and disgust situates from dislike to loathing. Both anger and disgust emotions can be express from various aspects such as food, service, price, dining environment, and even restaurant decoration. However, those emotions are not preferred by readers and cannot stimulate appetence for people to retweet. Sadness situates from gloominess to grief, and surprise situates from uncertainty to

المنارة للاستشارات
www.manaraa.com

amazement. Either sadness or surprise is a strong personal feeling during the dining. Emotions of sadness and surprise are hard to be felt by other people who do not have the same experience and do not involve in a specific context. Therefore, sadness emotion also does not have any addition to retweeting behavior. As to the joy, although it belongs to positive sentiment, it has a negative effect on the retweet number. Joy situates from serenity to ecstasy, due to such unnatural and discordant expression from tweets, in many cases, joy emotion normally expresses with exaggerated effects which may be probably regarded by customers as fabricated information posted by the farming robot or fake information. Therefore, joy also does not have a contribution to customer retweeting behavior.

Besides sentiments and emotions, LSM has a significant positive effect on the retweet number. Focusing on the verbal mimicry, LSM measured the functional words of textual information and predict changes of interests in the social-psychological aspect (Gonzales, Hancock, & Pennebaker, 2010) where a higher LSM refers to a lower social distance of communication (Chung & Pennebaker, 2007). The results in the current study demonstrate tweets with a low social distance of communication among people (high LSM score) contribute to the customer retweeting behavior, because the communication barriers among those people are less.

On the other hand, the analysis based on the timeline accordingly reflects the customer emotion fluctuation in six different time periods of a day. Based on the result, positive sentiment reaches a peak at 10:00am - 1:00pm. Associating with general dining behavior, this period of time usually refers to the lunchtime around noon. The phenomena can be well explained due to the lunch break around noon, people can have a respite from their work environment, so the positive emotions reach the peak. And the results are consistent with the analysis conducted by

المنارة للاستشارات
www.manaraa.com

Stone et al. (2006). However, during the time period from 2:00am - 5:00am, the negative sentiment expressed in customer tweets reaches the peak. Discussed by Wang et al. (2013), although people may have a bad feeling at night no matter related to their work or life or not, it would lead to a higher level of the negative sentiment that they feel in the next morning. For example, if people are over considering his or her performance in office or service quality at the restaurant at night, it could increase the bad mood in the next morning. Therefore, more negative emotions expressed in the early morning from 2:00am - 5:00am can be fully explained with this consideration. At the same time, combining the result of the radar chart, more detailed interpretations based on the timeline analysis can be discovered from each one of those eight separated emotional dimensions.

The geostatistical analysis applied in the current study mainly focused on investigations of the influence of two sentiments polar based on the geographic distance among the city, state, and nationwide levels. According to the perdition results, the positive sentiment indicates both with a higher level of geographic influence and a higher level of the spread than the negative sentiment. And the results are consistent with the negative binomial regression where positive sentiment can contribute to the retweet number, but the negative sentiment would inhibit customer retweeting behavior. Geostatistical analysis in the current study acts as a complementary role which further demonstrates the intuitive presented the diffusion of customer emotion and sentiments on the map according to the geographic influence.

المنارة للاستشارات

www.manaraa.com

CHAPTER 6. IMPLICATIONS

6.1. Academic implications

In the current study, the main variables of interests were developed based on the ELM theory, thus the results contribute to the ELM theory from many aspects as well. As mentioned above, the peripheral route of ELM can normally dominate the persuasion process when customers read or review online UGC, because this approach does not require much effort in information processing (Lumen, 2019). Although the research topics created based on the ELM in the hospitality industry are not scarcity and comprised of board study areas including customer behavior intention (e.g. Cheng & Loi, 2014; Hur et al., 2017), customer perception (e.g. Hu, 2012; Hussain et al., 2017), loyalty and satisfaction (e.g. Levy & Duverger, 2010; Yoo et al., 2017) according to Table 22 in Section 2.4, limited study spotlighted the customer retweeting or reposting behavior in the online social media and micro-blog context. Meanwhile, different from the surveys and regression methods that were widely conducted in the traditional investigation process, the current study innovatively integrated the big-data concept and incorporated text mining methods, and then explored some latent determinants based on the customer UGC information. Rather than only inspecting some superficial determinants such as retweeting numbers or the tweet length, investigation of the current study discovered sentiments, emotions, LSM that are embedded in each tweet provided by the customer, and identified the effects of those latent determinants on customer retweeting behavior. Therefore, in the current study, several aspects of theoretical implications were identified of ELM not only from the new variables' exploration but also from the innovative analysis processes.

Second, the current study also conducted extensive analyses on sentiments and emotions, and made significant contributions to the theory of Plutchik's emotional framework or emotion

المنارة للاستشارات

www.manaraa.com

wheel. As mentioned before, based on the viewpoint of Plutchik (1994) there are two general sentiments (i.e., positive and negative) of people's emotion feelings. However, under the domain of each sentiment, there are eight emotional dimensions identified from the social psychological perspective, which are further classified into four opposite pairs including sadness-joy, fear-anger, disgust-trust, and surprise-anticipation, and consist the whole emotion wheel from these eight separated dimensions referring to the emotion wheel in Figure 1. Introduced in Section 2.6, although Plutchik's emotional framework was one of the most established theories targeting the social science context, and some professional research organization such as National Research Council Canada (Mohammad & Turney, 2013) already developed extensive studies and even created an EmoLex dictionary to help future research, few studies neither implicated the EmoLex dictionary nor conducted the text mining process based on Plutchik's emotional framework. Therefore, on the one hand, the results in the current study provided contributions and further implications in the research method of emotional studies in the hospitality industry. On the other hand, it theoretically contributed to the study of Plutchik's emotional framework by exploring potential variables of interests and understanding customer behavioral intentions even in a boarder social science field.

Besides, the current focusing on the tweeting and retweeting behavior provided a new research perspective on the Technology Acceptance Model (TAM). Although according to the original model provided in Figure 19, perceived usefulness and perceived ease of use were identified as two determinates influencing customer's attitude and behavioral intentions (Davis, Bagozzi & Warshaw, 1989), a considerable of studies have made some improvements and created some extended TAM to test other variables besides the original two (Lee, Kozar, & Larsen, 2003). For example, Pavlou (2003) developed a research topic around customer

المنارة للاستشارات
www.manaraa.com

acceptance of electronic commerce and explored a new variable of perceived risk besides perceived usefulness and perceived ease of use. Therefore, taking advantages of the results in the current study, it provides theoretical implications for further research to investigate customer behavioral intention with new directions through combining the two sentiments, eight emotions, and LSM customers perceived as well. At the same time, referring to the regression model showing in Section 3.6, the current study only targeted the dependent variables of interests of behavioral intention (customer retweeting behavior) which skipped the middle process of customer attitude formation. Therefore, the future studies also can be developed on an integrated study process based on the whole framework of the TAM theory.

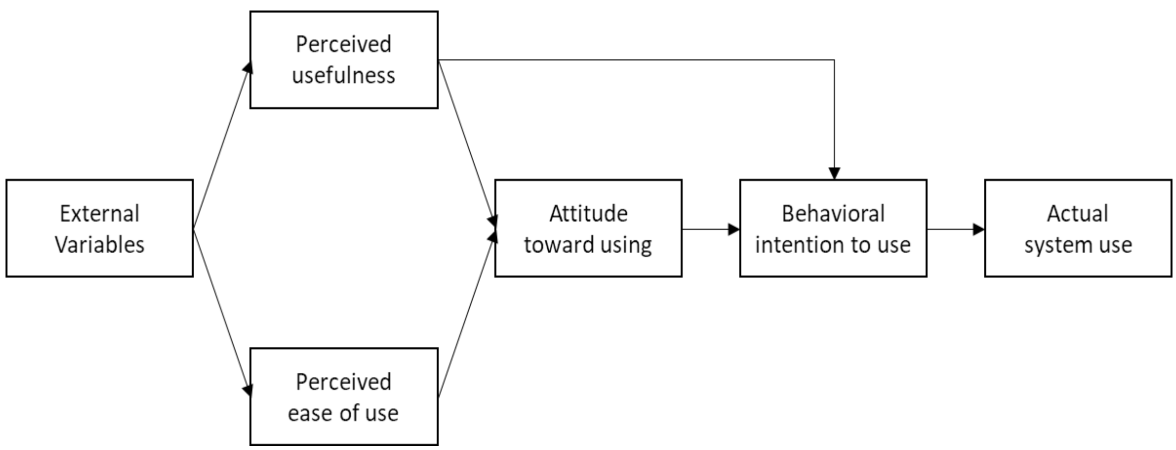

Figure 19. The Technology Acceptance Model

Last but not least, the current study also contributed to the literature of geostatistical analysis in the hospitality field. As mentioned before, different from the spatial analysis developed based on the areal data with human setting boundaries such as zip code, county, or city, geostatistical analysis required data with continuous effects without the constraint of the boundaries (Bivand et al., 2008, p263). And the expression of sentiments or emotions embedded with textual tweets content can have continuous effects on the neighboring observations. In

المنارة للاستشارات
www.manaraa.com

addition, the timeline analysis conducted in the current study investigated the features of emotions and sentiments from another perspective. With a combination of both geographic and time perspectives analyses, the current study was regarded as a pioneer exploration of the sentiment and emotion diffusions not only from a geographic aspect at a nationwide level to a narrow-downed city level, but also from a time aspect with six different time sequences in every day.

6.2. Practical implications

Results of the current study also provided many practical implications for individuals, hospitality business providers, social media platforms, and the government. First, as to the individual, many people nowadays want to become the social media influencer on the Internet, and there is also a group of people called "Internet celebrity" that has become a cultural phenomenon existing in the young generation (Juntiwasarakij, 2018). The Internet celebrity normally is followed by millions of Internet followers, and they are struggling for being "liked" "shared", or "reposted" in online social media platform (Juntiwasarakij, 2018), because they actually make their livings based on their Internet influence (Hou, 2013). Therefore, understanding the emotions and the functional words based on the results in the current study can be extremely useful when they would like to post some information online and generate a large volume of likes and reposts. At the same time, measuring nonconscious verbal matching (LSM) in their words, can facilitate smooth interpersonal interaction, and contribute to maintaining the relationship with followers as well. (Ireland et al., 2011). And associating with the emotion diffusion maps in Section 4.4, through a carefully carving and polishing process according to both of two sentiments, every individual emotional dimension, and each of nine functional word categories, can significantly contribute to the diffusion of the information, enlarge the extent of

www.manaraa.com

influence, and increase the number of followers. After that, not only the social influence on the target audiences, but also the business and marketing results can both increase by a decent degree.

In the same logic, businesses and organizations in the hospitality industry also can take advantage of the results in the current study when either doing marketing or public relationship. As mentioned in Section 2.5, studies have already discussed the emotions or sentiments in the hospitality industry to some level, however, the current study provided a holistic analysis of two sentiments, eight emotions, and nine functional word categories within a context social media and micro-blog platform of Twitter. Compared with previous literature, the current study not only opens a new window in the hospitality industry to explore the emotions from customers' tweets, but also provides foresight to the businesses and organizations when doing the marketing on Twitter. If hospitality businesses try to refine the marketing information presented to the customers according to the analysis in the current study, not surprisingly, it can contribute to accomplish better advertising results, formulate competitive advantage, and obtain more potential customers in the industry. In the meantime, businesses also can particularly focus the emotions in each timeline based on the results, and develop some specific strategies targeting a certain group of customers.

On the other hand, text mining and regression results in the current analysis also demonstrate a new business opportunity for the social media platform providers. Nowadays, people have already changed their life study, and the Internet and online social media platforms have become one of the most primary sources that people maintain the connections with externals (Wan, Tang, & Kim, 2019). As mentioned above, businesses prefer using online social media platforms as an important marketing channel, therefore, online social media platform

www.manaraa.com

providers also can be applied the horizontal integration strategy and provide with a diversity of services to the businesses and organizations related to their products and services. For example, restaurant review website www.yelp.com has developed the customer review recommendation system with a purpose to best reflect the opinions of the Yelp community (Yelp, 2020). Therefore, with this concern, the results in the current study can guide online social media providers to create some new functions based on the UGC information and textual content such as fake information deleting, click farming, and robot review detection, so as to satisfy the diverse needs and wants of businesses.

In addition, the current study also makes a contribution to the local government and destination management organizations (DMO) at tourist destinations. Analyzing the textual information provided by people or tourists on the online social media, it helps both local government and DMOs to have a well understanding about the topics along with emotions expressed by its tourists, so that local government and DMOs can do a better job in promoting tourist destinations and building up positive destination images to the people's minds. Besides, through the concerning of online social media, once there is a risk event and emergency condition, local government and DMOs can get real-time access to the opinions and attitudes and have a careful analysis on their sentiments and emotions, so as to provide an efficient crisis handing for maintaining public relation on time and keep the negative influences as minimal as possible.

المنارة للاستشارات
www.manaraa.com

CHAPTER 7. LIMITATION

The current study has several limitations. First, online UGC information still has many aspects that can be explored through the text mining method. In the current study, although it provided a holistic textual analysis on emotional dimensions and functional words, future studies can follow this logic and explore some potential variables such as cognition, social relationship, and time concept which can be generated from not only UGC contents, but also from the customer profile information.

Second, as to the geostatistical analysis, the current study only examined the relationship and analyzed the influence of two polar sentiments, future studies can also delve into eight different emotional dimensions, and then reveal effects of those sub-emotions on the geographic perspective. Meanwhile, due to the limitation of sample size, the current study only stopped at the mapping estimation, but did not have the statistical prediction and modeling processes. However, if narrowing down to a specific area of interests such as a city or a state, future studies can move a further step by using Kriging methods to get the geostatistical modeling results of both sentiments and emotions, and even other potential variables based on the geographic influence of distance.

Moreover, the current study investigated the effect of both sentiments and emotions during the six-period timeline and identified significant differences between sentiments and among emotions. Future studies also can go with this flow and explore the discrepancies among variables of interests on the weekly period or monthly period, so as to get a comprehensive understanding of the influence on the timeline perspective.

www.manaraa.com

CHAPTER 8. CONCLUSION

With the emergence of social media, online UGC has become one of the most important sources for information exchanging in all aspects of people's daily life. Considering the effect of online UGC, it not only can be regarded as a valuable opinion bank for business providers, but also enables a new online lifestyle for people as well as most of the potential customers of restaurants. Furthermore, after potential customers perceive information contained in the online UGC no matter in different formats, their feeling, attitude, and post behavior intention would have corresponding changes and adjustments as well.

Stimulated in this concern, the current study explored the textual content of Twitter, one of the largest micro-blog businesses around the world. With the purpose of having an exhaustive examination on the effects of different formats of information including the sentiments, emotions, and LSM (functional words) contained in customer posted tweets, a series of statistical analyses including Pearson correlation, negative binomial regression, t-test, and ANOVA and visualization processes including geographic mapping, line chart, and radar chart were conducted to an all-year-around dataset collected from Twitter with around one million pieces of tweet information.

The results of the current study ideally explained hypotheses propositions developed based on those three aspects of online UGC contents mentioned above, and all independent variables of interests including two sentiments, eight emotional dimensions, and the total LSM score for nine components of functional words were indicated having significant relationships with dependent variables of customer retweeting intention. In the specific cases, indicating from the first regression model, positive sentiment had a positive effect on retweeting behavior while negative sentiment had a negative effect on retweeting behavior. Revealing in the second

www.manaraa.com

regression model, Emotional dimensions of anger, disgust, joy, sadness, and surprise had negative effects on retweeting behavior, while anticipation, fear, and trust had positive effects on retweeting behavior. The total LSM score of nine categories of functional words had a positive effect on retweeting behavior as well. Besides, based on the timeline analysis, not only two sentiments but also eight emotions were identified with significant differences from a six-period timeline of one day. At the same time, focusing on both positive and negative sentiments, the geographic relationships were also identified separately at the city level, state level, and nationwide level based on geographic coordinates provided along with tweets.

Discussion interpreted the significance of the most valuable findings of the current study from threefold. First, focusing on the negative binomial regression results in two models, each of those two sentiments and eight emotional dimensions, no matter with positive or negative effects on the customer retweeting behavior, were fully discussed and well explained. Second, as to the results of language style matching, the reason why a low social distance of communication among people can increase the retweet number was also explained. Besides, based on the timeline and geostatistical analysis, the current study also provided new insights on why there were time differences and geographic differences of both sentiments and emotions that were expressed embedded with customer tweet contents.

Implications were also provided in the current study, in which some important insights were suggested to both academia and industry based on the detailed analysis results and findings. As to the academic implications, focusing on the text mining results from sentiment, emotions, and LSM, the current study not only contributes to the theory of ELM, but also extends the analysis of Plutchik's emotional framework to the social media platform in the hospitality industry. Focusing on the information diffusion process, it contributes to the TAM on analyzing

www.manaraa.com

customer retweeting behavior. At the same time, both timeline analysis and geostatistical analysis are also associated with customer sentiments and emotions with an innovative approach, which can be regarded as one of the pioneer studies in the hospitality field. As to the practical implications, results and findings in the current study made great contributions among individuals, hospitality businesses and originations, DMOs, and even local government at a certain tourist destination. It not only gives suggestions to someone who wants to become an Internet celebrity, but also provides business solutions to the hospitality organizations. It contributes to the process of option management for tourism DMOs as well as assists local government to come up with an efficient crisis public relation and minimize bad influences from negative emotions as much as possible.

Several limitations were added at the end of the current study from three aspects related to the online UCG, geostatistical analysis, and timeline analysis. Future studies could develop potential research topics from exploring more determinates from online UGC, investigating geographic relationship and estimation models by using Kriging methods, and testing the discrepancies among variables of interests on timeline form the different aspects such as weekly period or monthly period.

المنارة للاستشارات
www.manaraa.com

REFERENCES

Abbar, S., Mejova, Y., & Weber, I. (2015). You tweet what you eat: Studying food consumption through twitter. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3197-3206). New York, NY: ACM.

Adam, I., & Mensah, E. A. (2014). Perceived spatial agglomeration effects and hotel location choice. *Anatolia*, *25*(1), 49-60.

Adeborna, E., & Siau, K. (2014). An Approach to Sentiment Analysis-the Case of Airline Quality Rating. *Proceedings of the 2014 Pacific Asia Conference on Information Systems* (p. 363). Chengdu, China: PACIS.

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment analysis of twitter data. *Proceedings of the Workshop on Language in Social Media (LSM 2011)* (pp. 30-38). Stroudsburg, PA: LSM.

Antoniadis, K., Zafiropoulos, K., & Vrana, V. (2015). Communities of followers in tourism twitter accounts of European countries. *European Journal of Tourism, Hospitality and Recreation*, *6*(1), 11-26.

Azpurua, M. A., & Ramos, K. D. (2010). A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude. *Progress in electromagnetics research*, *14*, 135-145.

Balaguer, J., & Pernías, J. C. (2013). Relationship between spatial agglomeration and hotel prices. Evidence from business and tourism consumers. *Tourism Management*, *36*, 391-400.

Bifet, A., & Frank, E. (2010). Sentiment knowledge discovery in twitter streaming data. *Proceedings of International conference on discovery science* (pp. 1-15). Heidelberg, Germany: Springer.

Bivand, R. S., Pebesma, E. J., Gomez-Rubio, V., & Pebesma, E. J. (2008). *Applied spatial data analysis with R* (2nd ed.). New York, NY: Springer.

Bokunewicz, J. F., & Shulman, J. (2017). Influencer identification in Twitter networks of destination marketing organizations. *Journal of Hospitality and Tourism Technology*, *8*(2), 205-219.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, *2*(1), 1-8.

Boyd, D., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. *Proceedings of 2010 43rd Hawaii International Conference on System Sciences* (pp. 1-10). Washington, DC: IEEE.

المنارة للاستشارات

www.manaraa.com

Brandt, T., Bendler, J., & Neumann, D. (2017). Social media analytics and value creation in urban smart tourism ecosystems. *Information & Management*, *54*(6), 703-713.

Brunner-Sperdin, A., Peters, M., & Strobl, A. (2012). It is all about the emotional state: Managing tourists' experiences. *International Journal of Hospitality Management*, *31*(1), 23-30.

Chang, K. C. (2016). Effect of servicescape on customer behavioral intentions: Moderating roles of service climate and employee engagement. *International Journal of Hospitality Management*, *53*, 116-128.

Chung, C., & Pennebaker, J. W. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social communication* (pp. 343-359). New York, NY: Psychology Press.

Claster, W. B., Cooper, M., & Sallis, P. (2010). Thailand-Tourism and conflict: Modeling sentiment from Twitter tweets using naïve Bayes and unsupervised artificial neural nets. *Proceedings of 2010 Second International Conference on Computational Intelligence, Modelling and Simulation* (pp. 89-94). Bali, Indonesia: IEEE.

Claster, W., Pardo, P., Cooper, M., & Tajeddini, K. (2013). Tourism, travel and tweets: algorithmic text analysis methodologies in tourism. *Middle East Journal of Management*, *1*(1), 81-99.

da Silva, F. M., de Carvalho Alves, M., de Lima Bueno, R., & da Costa, P. A. N. (2012). Geostatistical analysis of fruit yield and detachment force in coffee. *Precision Agriculture*, *13*(1), 76-89.

Da Silva, N. F., Hruschka, E. R., & Hruschka Jr, E. R. (2014). Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, *66*, 170-179.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science*, *35*(8), 982-1003.

De Albornoz, J. C., Plaza, L., Gervás, P., & Díaz, A. (2011). A joint model of feature mining and sentiment analysis for product review rating. *Proceedings of European conference on information retrieval* (pp. 55-66). Heidelberg, Germany: Springer.

De Rosario, A. H., Rodríguez, M. D. M. G., & Pérez, M. D. C. C. (2013). Development of social media and web 2.0 in the top hotel chains. *Tourism & Management Studies*, *9*(1), 13-19.

Deng, W. J., Yeh, M. L., & Sung, M. L. (2013). A customer satisfaction index model for international tourist hotels: Integrating consumption emotions into the American Customer Satisfaction Index. *International Journal of Hospitality Management*, *35*, 133-140.

Diao, Q., & Jiang, J. (2014). Recurrent Chinese restaurant process with a duration-based discount for event identification from twitter. *Proceedings of the 2014 SIAM International Conference on Data Mining* (pp. 388-397). Philadelphia, PA: SDM.

www.manaraa.com

Duffett, R. G. (2015). Facebook advertising's influence on intention-to-purchase and purchase amongst Millennials. *Internet Research. 25*(4), 498-526.

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169-200.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, *52*, 498-506.

Felbermayr, A., & Nanopoulos, A. (2016). The role of emotions for the perceived usefulness in online customer reviews. *Journal of Interactive Marketing*, *36*, 60-76.

Francisco, V., & Gervás, P. (2006). Automated mark up of affective information in English texts. *International Conference on Text, Speech and Dialogue* (pp. 375-382). Springer, Berlin, Heidelberg.

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological bulletin*, *118*(3), 392.

Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences*. Independence, KY: Thomson Brooks/Cole Publishing Co.

Gill, A. J., French, R. M., Gergle, D., & Oberlander, J. (2008). The language of emotion in short blog texts. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 299-302). San Diego, CA: ACM.

Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language style matching as a predictor of social dynamics in small groups. *Communication Research, 37*(1), 3-19.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation.* New York, NY: Oxford University Press.

Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of hydrology*, *228*(1-2), 113-129.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice hall.

Han, H., & Jeong, C. (2013). Multi-dimensions of patrons' emotional experiences in upscale restaurants and their role in loyalty formation: Emotion scale improvement. *International Journal of Hospitality Management*, *32*, 59-70.

Hay, B. (2010). Twitter Twitter-but Who Is Listening?: A Review of the Current and Potential Use of Twittering as a Tourism Marketing Tool. *Proceedings of CAUTHE 2010: Tourism and Hospitality: Challenge the Limits*, (p. 561). Tasmania, Australia: CAUTHE.

www.manaraa.com

Hays, S., Page, S. J., & Buhalis, D. (2013). Social media as a destination marketing tool: its use by national tourism organisations. *Current issues in Tourism*, *16*(3), 211-239.

Hengl, T. (2009). *A practical guide to geostatistical mapping* (2nd ed.). Luxembourg: Office for Official Publications of the European Communities.

Henningsen, M. L. M., Henningsen, D. D., Cruz, M. G., & Morrill, J. (2003). Social influence in groups: A comparative application of relational framing theory and the elaboration likelihood model of persuasion. *Communication Monographs*, *70*(3), 175-197.

Hou, M. (2013). The semiotics of Internet celebrity: Gangnam Style case. *Tilburg Papers in Culture Studies*. Tilburg, Netherlands: Tilburg University.

Hwang, Y., & Mattila, A. S. (2019). Feeling left out and losing control: The interactive effect of social exclusion and gender on brand attitude. *International Journal of Hospitality Management*, *77*, 303-310.

Hyun, S. S., Kim, W., & Lee, M. J. (2011). The impact of advertising on patrons' emotional responses, perceived value, and behavioral intentions in the chain restaurant industry: The moderating role of advertising-induced arousal. *International Journal of Hospitality Management*, *30*(3), 689-700.

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological science*, *22*(1), 39-44.

Ireland, M.E., & Pennebaker, J.W. (2010). Language style matching in writing: Synchrony in essays, correspondence, and poetry. Journal of Personality and Social Psychology, 99, 549–571.

Jang, S. S., Ha, J., & Park, K. (2012). Effects of ethnic authenticity: Investigating Korean restaurant customers in the US. *International Journal of Hospitality Management*, *31*(3), 990-1003.

Jani, D., & Han, H. (2015). Influence of environmental stimuli on hotel customer emotional loyalty response: Testing the moderating effect of the big five personality factors. *International Journal of Hospitality Management*, *44*, 48-57.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, *60*(11), 2169-2188.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* (pp. 56-65). San Jose, CA: ACM.

Jiang, L., & Erdem, M. (2017). Twitter-marketing in multi-unit restaurants: is it a viable marketing tool?. *Journal of foodservice business research*, *20*(5), 568-578.

المنارة للاستشارات
www.manaraa.com

Joyce, E., & Kraut, R. E. (2006). Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication*, *11*(3), 723-747.

Jung, H. S., & Yoon, H. H. (2011). The effects of nonverbal communication of employees in the family restaurant upon customers' emotional responses and customer satisfaction. *International Journal of Hospitality Management*, *30*(3), 542-550.

Juntiwasarakij, S. (2018). Framing emerging behaviors influenced by internet celebrity. *Kasetsart Journal of Social Sciences*, *39*(3), 550-555.

Kemp, S. (2019). *DIGITAL 2019.* Hootsuite. https://p.widencdn.net/kqy7ii/Digital2019-Report-en

Kim, D., & Jang, S. S. (2015). Cognitive decline and emotional regulation of senior consumers. *International journal of hospitality management*, *44*, 111-119.

Kim, J. H., & Jang, S. S. (2014). A scenario-based experiment and a field study: A comparative examination for service failure and recovery. *International Journal of Hospitality Management*, *41*, 125-132.

Kim, J. H., Youn, H., & Rao, Y. (2017). Customer responses to food-related attributes in ethnic restaurants. *International journal of hospitality management*, *61*, 129-139.

Kim, S., Bak, J. Y., & Oh, A. H. (2012). " Discovering emotion influence patterns in online social network conversations" by Suin Kim, JinYeong Bak, and Alice Oh, with Ching-man Au Yeung as coordinator. *ACM SIGWEB Newsletter*, *3*, 1-6.

Kim, S., Bak, J. Y., & Oh, A. H. (2012). Do you feel what I feel? social aspects of emotions in twitter conversations. In *Sixth International AAAI Conference on Weblogs and Social Media*.

Kulshrestha, J., Kooti, F., Nikravesh, A., & Gummadi, K. P. (2012). Geographic dissection of the twitter network. *Proceedings of Sixth International AAAI Conference on Weblogs and Social Media*. Dublin, Ireland: AAAI.

Kwon, N., & Choi, Y. G. (2014). The relationships among coffeehouse's physical environment, self-congruity, positive emotion, and revisit intentions. *Culinary Science & Hospitality Research*, *20*(5), 111-118.

Lee, W., Xiong, L., & Hu, C. (2012). The effect of Facebook users' arousal and valence on intention to go to the festival: Applying an extension of the technology acceptance model. *International Journal of Hospitality Management*, *31*(3), 819-827.

Lee, Y., Kozar, K. A., & Larsen, K. R. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for information systems*, *12*(1), 50.

المنارة للاستشارات
www.manaraa.com

Lin, I. Y., & Worthley, R. (2012). Servicescape moderation on personality traits, emotions, satisfaction, and behaviors. *International Journal of Hospitality Management*, *31*(1), 31-42.

Liu, S., Law, R., Rong, J., Li, G., & Hall, J. (2013). Analyzing changes in hotel customers' expectations by trip mode. *International Journal of Hospitality Management*, *34*, 359-371.

Loureiro, S. M. C. (2014). The role of the rural tourism experience economy in place attachment and behavioral intentions. *International Journal of Hospitality Management*, *40*, 1-9.

Loureiro, S. M. C., Almeida, M., & Rita, P. (2013). The effect of atmospheric cues and involvement on pleasure and relaxation: The spa hotel context. *International Journal of Hospitality Management*, *35*, 35-43.

Luo, Y., & Tang, R. L. (2019). Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *International Journal of Hospitality Management*, *80*, 144-154.

Luo, Z., Osborne, M., Tang, J., & Wang, T. (2013). Who will retweet me? Finding retweeters in Twitter. *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval* (pp. 869-872). New York NY: ACM.

Lwin, M., & Phau, I. (2012). Exploring the effectiveness of existential guilt appeal: durable products. *Proceedings of the 41st EMAC conference 2012*. Lisbon, Portugal: EMAC.

Macskassy, S. A., & Michelson, M. (2011). Why do people retweet? anti-homophily wins the day!. *Fifth International AAAI Conference on Weblogs and Social Media* (pp. 209-216). Barcelona, Spain: AAAI.

Meng, B., & Choi, K. (2017). Theme restaurants' servicescape in developing quality of life: The moderating effect of perceived authenticity. *International Journal of Hospitality Management*, *65*, 89-99.

Miao, L., Mattila, A. S., & Mount, D. (2011). Other consumers in service encounters: a script theoretical perspective. *International Journal of Hospitality Management*, *30*(4), 933-941.

Min, H. K., & Kim, H. J. (2019). When service failure is interpreted as discrimination: Emotion, power, and voice. *International Journal of Hospitality Management*, *82*, 59-67.

Minazzi, R. (2015). *Social media marketing in tourism and hospitality*. New York, NY: Springer.

Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Cambridge, MA: Academic Press.

المنارة للاستشارات

www.manaraa.com

Misopoulos, F., Mitic, M., Kapoulas, A., & Karapiperis, C. (2014). Uncovering customer service experiences with Twitter: the case of airline industry. *Management Decision*, *52*(4), 705-723.

Mohammad, S. M., & Turney, P. D. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26-34). Los Angeles, CA: ACL.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436-465.

Molyneux, L. (2015). What journalists retweet: Opinion, humor, and brand development on Twitter. *Journalism*, *16*(7), 920-935.

Morosan, C., & DeFranco, A. (2015). Disclosing personal information via hotel apps: A privacy calculus perspective. *International Journal of Hospitality Management*, *47*, 120-130.

Mostafa, M. M. (2013). An emotional polarity analysis of consumers' airline service tweets. *Social Network Analysis and Mining*, *3*(3), 635-649.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, *41*(5), 673-690.

Park, S. B., Jang, J., & Ok, C. M. (2016). Analyzing Twitter to explore perceptions of Asian restaurants. *Journal of Hospitality and Tourism Technology*, *7*(4), 405-422.

Park, S. B., Ok, C. M., & Chae, B. K. (2016). Using twitter data for cruise tourism marketing and research. *Journal of Travel & Tourism Marketing*, *33*(6), 885-898.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International journal of electronic commerce*, *7*(3), 101-134.

Peng, C., Yin, D., Wei, C., & Zhang, H. (2014). How and when review length and emotional intensity influence review helpfulness: empirical evidence from epinions.com. *proceedings of Thirty Fifth International Conference on Information Systems* (pp. 1-16). Auckland, New Zealand: CIS.

Peng, N., Chen, A., & Hung, K. P. (2017). The effects of teppanyaki restaurant stimuli on diners' emotions and loyalty. *International Journal of Hospitality Management*, *60*, 1-12.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In Communication and persuasion. New York, NY: Springer.

www.manaraa.com

Petty, R. E., Barden, J., & Wheeler, S. (2009). The Elaboration Likelihood Model of persuasion: Developing health promotions for sustained behavioral change. In R. J. DiClemente, R. A. Crosby, & M. C. Kegler (Eds.), *Emerging theories in health promotion practice and research* (pp. 1-35). San Francisco, CA: Jossey-Bass.

Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management*, *55*, 16-24.

Phillips-Wren, G. E., & Hoskisson, A. (2014). Decision Support with Big Data: A Case Study in the Hospitality Industry. *Frontiers in Artificial Intelligence and Applications*, *261*, 401-413.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169-226.

Plutchik, R. (1994). *The psychology and biology of emotion*. New York, NY: HarperCollins College Publishers.

Restaurant Business. (2018). *TOP 250: THE RANKING.* Winsight. https://www.restaurantbusinessonline.com/top-500-chains?year=2018#data-table

Ruiz-Alba, J. L., Nazarian, A., Rodríguez-Molina, M. A., & Andreu, L. (2019). Museum visitors' heterogeneity and experience processing. *International Journal of Hospitality Management*, *78*, 131-141.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. *Proceedings of the 19th international conference on World wide web* (pp. 851-860). New York, NY: ACM.

Sevin, E. (2013). Places going viral: Twitter usage patterns in destination marketing and place branding. *Journal of Place Management and development*, *6*(3), 227-239.

Shimada, K., Inoue, S., & Endo, T. (2012). On-site likelihood identification of tweets for tourism information analysis. *Proceedings of 2012 IIAI International Conference on Advanced Applied Informatics* (pp. 117-122). Fukuoka, Japan: IEEE.

Siamionava, K., Slevitch, L., & Tomas, S. R. (2018). Effects of spatial colors on guests' perceptions of a hotel room. *International Journal of Hospitality Management*, *70*, 85-94.

Singal, M. (2012). Effect of consumer sentiment on hospitality expenditures and stock returns. *International Journal of Hospitality Management*, *31*(2), 511-521.

Song, J., & Qu, H. (2017). The mediating role of consumption emotions. *International Journal of Hospitality Management*, *66*, 66-76.

Sotiriadis, M. D., & Van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research*, *13*(1), 103-124.

المنارة للاستشارات

www.manaraa.com

Stieglitz, S., & Dang-Xuan, L. (2012). Political communication and influence through microblogging—An empirical analysis of sentiment in Twitter messages and retweet behavior. In *2012 45th Hawaii International Conference on System Sciences* (pp. 3500-3509). Piscataway, NJ: IEEE.

Stone, A. A., Schwartz, J. E., Schkade, D., Schwarz, N., Krueger, A., & Kahneman, D. (2006). A population approach to the study of emotion: diurnal rhythms of a working day examined with the Day Reconstruction Method. *Emotion*, *6*(1), 139.

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *2010 IEEE Second International Conference on Social Computing* (pp. 177-184). Piscataway, NJ: IEEE.

Sukhu, A., Choi, H., Bujisic, M., & Bilgihan, A. (2019). Satisfaction and positive emotions: A comparison of the influence of hotel guests' beliefs and attitudes on their satisfaction and emotions. *International Journal of Hospitality Management*, *77*, 51-63.

Tang, L. R., Kim, J., & Wang, X. (2019). Estimating spatial effects on peer-to-peer accommodation prices: Towards an innovative hedonic model approach. *International Journal of Hospitality Management*, *81*, 43-53.

Tantanatewin, W., & Inkarojrit, V. (2018). The influence of emotional response to interior color on restaurant entry decision. *International Journal of Hospitality Management*, *69*, 124-131.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406-418.

Tobler, W. (2004). On the first law of geography: A reply. *Annals of the Association of American Geographers*, *94*(2), 304-310.

Torres, E. N., Wei, W., Hua, N., & Chen, P. J. (2019). Customer emotions minute by minute: How guests experience different emotions within the same service environment. *International Journal of Hospitality Management*, *77*, 128-138.

Tsao, W. C., Hsieh, M. T., Shih, L. W., & Lin, T. M. (2015). Compliance with eWOM: The influence of hotel reviews on booking intention from the perspective of consumer conformity. *International Journal of Hospitality Management*, *46*, 99-111.

Vidal, L., Ares, G., Machín, L., & Jaeger, S. R. (2015). Using Twitter data for food-related consumer research: A case study on "what people say when tweeting about different eating situations". *Food Quality and Preference*, *45*, 58-69.

Vo, T. T., Xiao, X., & Ho, S. Y. (2017). How does corporate social responsibility engagement influence word of mouth on Twitter? Evidence from the airline industry. *Journal of Business Ethics*, 1-18.

المنارة للاستشارات
www.manaraa.com

Wang, M., Liu, S., Liao, H., Gong, Y., Kammeyer-Mueller, J., & Shi, J. (2013). Can't get it out of my mind: Employee rumination after customer mistreatment and negative mood in the next morning. *Journal of Applied Psychology*, *98*(6), 989.

Wang, X., Tang, L. R., & Kim, E. (2019). More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness?. *International Journal of Hospitality Management*, *77*, 438-447.

Wang, Y. C. (2015). A study on the influence of electronic word of mouth and the image of gastronomy tourism on the intentions of tourists visiting Macau. *Turizam: međunarodni znanstveno-stručni časopis*, *63*(1), 67-80.

Weng, J., Lim, E. P., Jiang, J., & He, Q. (2010). Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 261-270). New York, NY: ACM.

Wu, L. L., Mattila, A. S., & Han, J. R. (2014). Territoriality revisited: Other customer's perspective. *International Journal of Hospitality Management*, *38*, 48-56.

Wu, L., Mattila, A. S., & Hanks, L. (2015). Investigating the impact of surprise rewards on consumer responses. *International Journal of Hospitality Management*, *50*, 27-35.

Xie, D., & Heung, V. C. (2012). The effects of brand relationship quality on responses to service failure of hotel consumers. *International Journal of Hospitality Management*, *31*(3), 735-744.

Xie, H. J., Miao, L., Kuo, P. J., & Lee, B. Y. (2011). Consumers' responses to ambivalent online hotel reviews: The role of perceived source credibility and pre-decisional disposition. *International Journal of Hospitality Management*, *30*(1), 178-183.

Yadav, M. L., & Roychoudhury, B. (2019). Effect of trip mode on opinion about hotel aspects: A social media analysis approach. *International Journal of Hospitality Management*, *80*, 155-165.

Yelp. (2020). *Recommended Reviews.* Yelp. https://www.yelp-support.com/Recommended_Reviews?l=en_US

Yin, G., Wei, L., Xu, W., & Chen, M. (2014). Exploring Heuristic cues for Consumer Perceptions of Online Reviews Helpfulness: the Case of Yelp.Com. *Proceedings of the 2014 Pacific Asia Conference on Information Systems (PACIS)* (pp. 52-64). Chengdu, China: PACIS.

Yoo, C. W., Goo, J., Huang, C. D., Nam, K., & Woo, M. (2017). Improving travel decision support satisfaction with smart tourism technologies: A framework of tourist elaboration likelihood and self-efficacy. *Technological Forecasting and Social Change*, *123*, 330-341.

المنارة للاستشارات

www.manaraa.com

Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism management perspectives*, *10*, 27-36.

Zhao, X. R., Wang, L., Guo, X., & Law, R. (2015). The influence of online reviews to online hotel booking intentions. *International Journal of Contemporary Hospitality Management*.

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, *76*, 111-121.

Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, *15*(4), 267.

المنارة للاستشارات

www.manaraa.com