

Chinese Online Reviews for Tourism Attractions: An Affective Computing System based on Data Mining

Yu Huang¹ Mu Zhang^{2*}

¹ School of Management, Jinan University, 601 Huangpu Avenue West, Guangzhou 510632, China

² GIS & Tourism Information Technology Laboratory, Shenzhen Tourism College, Jinan University, 6 Qiaochengdong Street, Shenzhen 518053, China

* Corresponding Author, zhangmu@163.com

Abstract:

Based on the excessive online reviews of tourism spot and related research status, this paper attempts to design an affective computing system to provide a comprehensive analysis of affective inclination toward multiple products/services in specific tourism spots, thus transforming complex and unstructured review texts into structured attitudinal information. The designing process consists of three parts: firstly, the author puts forward the framework of affective computing system of online tourism spot reviews based on modular design; secondly, build and modify a feature-oriented domain dictionary through association algorithm for feature mining; thirdly, conduct the mining of adverbs of degree, polarity and negative words grounded on the dependency under syntactic analysis; last of all, come up with the calculation method of online tourism spot reviews' affective inclination, construct and display the basic affective computing system.

Key words: tourism spot; online review; syntactic analysis; feature mining; affective computing

1 Introduction

1.1 Background and significance

With the rapid development of E-Commerce nowadays, online consumption has more or less integrated into a variety of industries. Tourism is no exception either. Currently, as much as 71% of individual tour and 36% of group tour are booked through online channel. Hotels, airlines and restaurants have also received more orders from the Internet. Meanwhile, the majority of tourism companies like scenic spots and transportation regard online sales platform such as OTA (Online Travel Agent) and independent booking agent as their largest distribution channel.

In addition, the rise of the idea "Web 2.0" and social network have encouraged consumers to give online feedbacks after their consumption, among which an important form is online review. By virtue of promptness, instantaneity and accessibility, it has become a type of word of mouth in digital age (Kaplan & Haenlein, 2010). Instead of official websites, potential consumers tend to obtain relatively objective online reviews from unofficial channels (Forman, et al., 2008), since they are more authentic and helpful for them to make purchase decisions (Li & Hitt, 2008). In tourism industry, where consumer satisfaction is greatly valued, the reviews play a more significant role in online sales practices, which accounts for the large amount of time many potential customers spend reading online reviews when making decisions (Zhu & Zhang, 2010).

In fact, online review is turning into a strong electronic type of word of mouth (Chen & Xie, 2008). Good reputation could bring generous price premium, therefore meeting company's need to pursue maximized value (Ye, et al., 2011). At present, the fierce competition in online sales market has driven companies to try all possible means to attract customers' attention. Despite that, online customers have remained their patience browsing online reviews, which indicates the magnificent part online review platforms play in promoting consumption and sharing information (Zhong et al., 2013).

However, the soaring number of online reviews has led to an increasing phenomenon of information overload and redundancy, thus made the core information and attitudes unavailable in short time. The mainstream solution to the problem is to screen out the reviews in the top ranking by usefulness vote, posting time and reviewer qualification, etc. But it sacrifices the completeness of online review contents. It is therefore of great importance to collect comprehensive reviews and sort out effective information (positive or negative inclination), and make it quickly available for both consumers and corporate managers.

1.2 Research status and trend

1.2.1 Online review for OTA

As the influence of online review on consumers and sellers got recognized, an increasing number of scholars from various fields have turned their focus to it, including those from tourism (Lee & Hu, 2005). Based on keyword searches (Due to limited researches that are fully relevant to online review of tourism spot, the paper chooses 'hotel, restaurant, spot, hospitality, travel and tourism' as the keywords to review the studies of online review in tourism), Science Direct and EBSCO database, this study makes an overview of representative researches in recent years, of which the areas can be classified into four groups as below.

(1) Online review and online purchase

The development of E-Commerce has greatly transformed consumers' pattern of obtaining information, making reference and purchase decision. Potential tourism consumers always hope to know beforehand: which spot, hotel, and restaurant are the best? That's when they already tend to get information from people with relevant experiences. In this sense, online review is important to both consumers and sellers. There emerged studies on the influences of online review on purchase intention, price and sales volume.

Scholars such as Mauri (2013) and Sparks (2011) pointed out the significant correlation between online reviews and purchase intention on tourist hotels. Miao (2011) found that high rating spots stimulate stronger revisit intentions. Similarly, Zhang (2013) proved the correlation between online reviews and tourism food consumption. Yacouel (2012) suggested that hotels with better reputation by online reviews can obtain higher price premium, which means larger profits. Besides, Ye (2011) discovered that 10% increase of the rating by online reviews would bring at least 5% rise in hotel and scenic spots' sales volume. Similar results were found by Lu (2013) in the restaurant studies.

(2) Customer satisfaction and online response

For the consumers who have completed consumption process, online review is the most common approach to express their feelings, both complaint and satisfaction. The collection and analysis of such information help corporate managers improve themselves (Pantelidis, 2010). As a matter of fact, it is always important for tourism enterprises to know customers' (dis)satisfaction, whatever type it belongs to.

At the same time, company's management of online response has also become more and more crucial. Positive online response could help company turn unsatisfied customers into loyal ones (Pantelidis, 2010). Effective response strategies toward online response, particularly negative ones, is useful for company to improve itself in the future.

(3) Motives behind online review posting/browsing/sharing

With the rise in online reviews' number and importance, researchers began to think about the major participants and their motives behind participation, who later introduced consumer psychology and behavior theory into the study so as to explore the trigger mechanism behind online review posting/browsing/sharing. Yoo & Gretzel (2008) believe that motive is a vision that company could improve itself and offer better experience for future customers. In their subsequent study, the two found that female visitors tend to help company improve service, while male visitors hope to warn future customers of the "traps" (Yoo & Gretzel, 2011). Oguta (2012) argues that existing rating, complaining emotion and low price are the motivational elements that inspire visitors to participate in online reviews. In their study on restaurants, Jeong and Jang (2011) found that delicious food and good manner of service are the key motives behind customers' participation. Kim (2011), however, points out that convenience, quality and risk reduction and social pressure are among the critical motives.

(4) Opinion mining and syntactic analysis

The soaring number of online reviews has led to an increasing phenomenon of information overload and redundancy. It is necessary to understand the emphasis, commendatory or derogatory meaning of them. Relevant research has naturally approached technical use of big data and data mining, evolving algorithm and programming. In contrast, human judging would take too much time and effort. Akehurst (2009) recommends the introducing of artificial intelligence and opinion mining technics.

There are overall two applications of opinion mining: value judgment and feature extraction. The former divides online reviews into the commendatory or derogatory sorts, which needs three mainstream auxiliary algorithms: Navie Bayes, SVM, N-gram Model. The last two prove better than the first one (Ye et al., 2009). In order to improve accuracy, Kang, Yoo & Han (2012) made certain modifications to such classifying technics. Feature extraction means extracting relevant information from online reviews according to syntactic and semantic paradigms.

To conclude, there are altogether four research topics regarding tourism online review: online review and online purchase, customer satisfaction and online response, motives behind online review posting/browsing/sharing, opinion mining and syntactic analysis. In terms of research subject, most people focus on tourism hotels, few paying attention to tourism spots.

1.2.2 Affective computing in tourism

Overseas studies on affective computing have received wide attention since 1990s. As of now, there exists a large quantity of remarkable studies that target the English texts.

During the past 10 years, the tourism boom has put affective computing under the spotlight. Tourism product can't be tried out, so tourists have to depend more on their subjective judgment, as opposed to objective measuring. Potential tourists have very limited knowledge on the spot they have never been, and they lack subjective accessibility to such information, which is crucial for them to make travel decisions.

As a result, relevant researches haven't gone through a long period at home. Classic ones include: Pan (2007) analyzed the complete evaluation of a tourism spot by affective computing; Ye and Zhang (2007) conducted automatic identification of Chinese texts' subjectivity; Ye (2009) employed SVM approach to classify

emotions toward tourism spots in English contexts and obtained good outcomes; same approach was used to classify relevant travelling questions (Zhang, 2009).

Based on previous research outcomes and current application status, this paper attempts to design a review mining system toward tourism spots. The procedure includes: system framework—review collection module—feature mining module—ffective computing module—systematic module integration. The design of all the modules will realize automatic collection and affective computing of online spot reviews and show consumers/platform runners/spot runners to help making decisions in their fields.

2 Basic definitions and system framework

2.1 Basic definitions

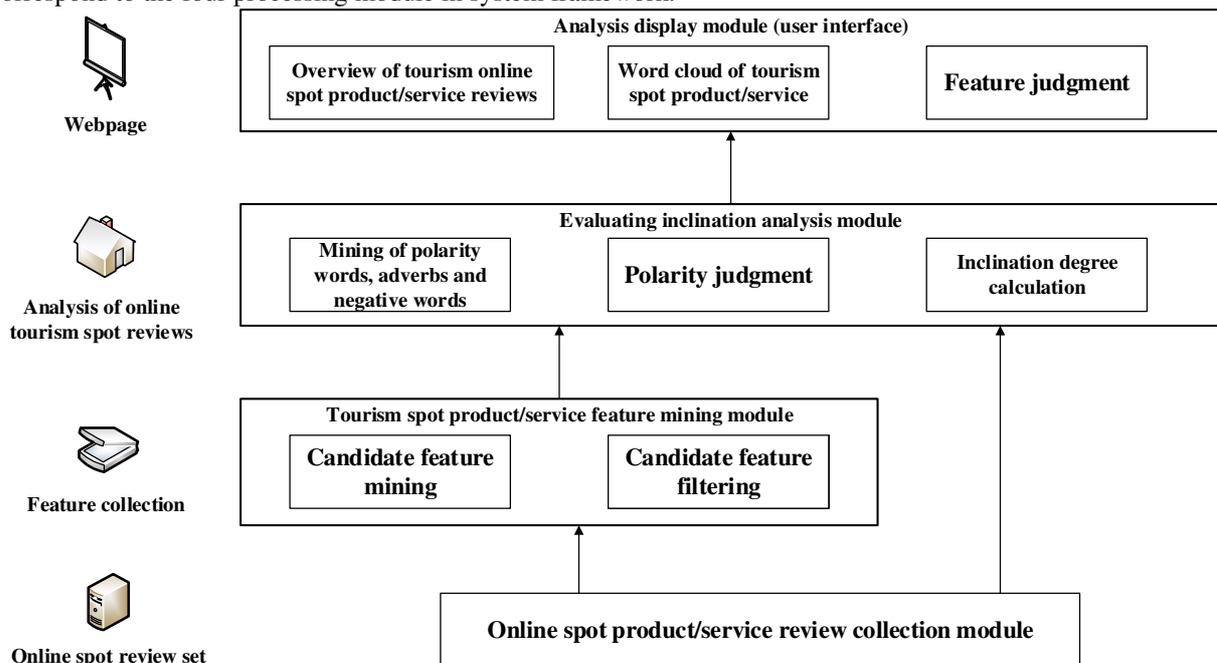
The primary task is to conduct feature mining and affective computing of online spot reviews. So we define the reviews as $R = \{r_1, r_2, \dots, r_m\}$ (m stands for the number of reviews toward a certain spot.)

$F(R) = \{f_1, f_2, \dots, f_n\}$ means the feature set within R (n stands for the number of features,)

We further define single feature review tuple as $O = (f', oword, dword, nword, isNDP)$. f' stands for feature word; $oword$ is polarity word; $dword$ being adverb of degree; $nword$ being negative word; $isNDP$ stands for the word order of negative word and adverb of degree. In fact, for each online review, the feature review tuple may be more than one. So the computing result of one review can be $\{o_1, o_2, \dots, o_j\}$ (j stands for the number of feature review tuple)

2.2 Function analysis-based system framework

From the perspective of function analysis, the affective computing system of online spot review mainly include four parts: online review collection—feature mining—ffective computing—result display, which correspond to the four processing module in system framework.



Graph 2-1 system framework diagram

As shown by graph 2-1:

Step 1: tourism product/service online review collection module. Collect and sort out automatically review data set R and provide data basis for subsequent feature mining and affective computing;

Step 2: product/service feature mining module. Use association algorithm to mine product/service feature based on collected review set R . As the review subject, product/service is the premise and main body for computing consumers' attitudes toward specific feature..

Step 3: affective computing module. Based on feature items, obtain polarity words and corresponding adverbs of degree and negative words via relevant algorithm, compute affective polarity and calculate by degrees.

Step 4: result display module. Organize computing result display according to users' browsing needs, including overall evaluation, main features, positive and negative reviews of different features. Allow users to interact with product/service features by permitting them to select those they are interested in and view the specific attitude and full review.

3 Design and application of key algorithms

This study mainly involves two algorithms: association rules based feature mining algorithm, including the mining of adverbs of degree, polarity and negative words, and affective computing based review inclination algorithm, including the calculation method of affective inclination degree.

3.1 Association rules based feature mining

We first obtain candidate feature sets through association algorithm, and screen out non-phrasal features by minimum independency support degree. Then we introduce self-defined field dictionary to supplement the mining of infrequent feature words so as to get more comprehensive and accurate product/service features. Meanwhile, semi-auto update of filed feature dictionary can be realized through the feedback from feature frequency mining, which will save a great deal of human work.

(1) Association-based mining of candidate feature sets

To begin with, we make word segmentation and part-of-speech tagging toward the reviews r in set R one by one. Then we rule out one-word nouns and put the rest into basic transaction $D = \{(TID, Noun(r_{TID}))\}$. Apriori algorithm will work out candidate feature frequency 1-itemset L_1 , frequency 2-itemset L_2 , etc.

(2) Filtering non-phrasal features

The mining algorithm of association rules based frequency item set doesn't take noun's position in review into consideration. So the two words in frequency 2-itemset L_2 may not be able to compose phrase. We try to filter noun phrases and obtain a new frequency 2-itemset L_2 .

(3) Filtering high frequency non-field features and adding low frequency features

Filter frequency 1-itemset in non-field feature dictionary from candidate feature set L_1 and add non-frequency 1-itemset in field feature dictionary to get L_1' .

(4) Filtering redundant features

A word's minimum independency support equals the support of frequency 1-itemset minus the absolute support of frequency 2-itemset (Itemset 1 includes the feature word; itemset 2 include the word). Set the threshold value of minimum independency support as 2, and filter the feature if it's lower than the value. Ultimately, we could obtain frequency 1-itemset L_1'' that meets the constraint of independency support.

(5) Updating field (non)feature dictionary

Dictionary update is realized through feedbacks. Users can choose themselves whether to judge the unlogged high frequency word list: if they decide the word is a field feature, the weight of its dependency on field feature dictionary will increase; otherwise the weight on non-field feature dictionary will increase. The dependency of the candidate feature is decided by the calculation of accumulated user feedbacks. The process is not compulsory. The cyclic iteration of seed dictionary could effectively enhance the mining function of similar product reviews.

After the candidate feature mining, filtering and supplementing non-frequent itemsets, we finally obtain the product feature set: $F(R) = L_2' \cup L_1' \cup L_1''$.

3.2 Syntactic analysis based lexical mining

On the basis of feature set F , this part traverses product review R by means of dictionary and syntactic analysis. HowNet affective lexicon and dependency information will be used to explore the relationship between polarity words and product features, combined with the sentiment polarity of the words outside lexical similarity computing network and synonym dictionary. It will also take into consideration the intensity gap between adverbs of degree, semantic gaps caused by them and negative words' co-occurrence order. Finally the degree of affective inclination will be calculated step by step. The application of lexical similarity computing and synonym dictionary would improve accuracy of affective inclination judgment.

3.2.1 Dependency based oword mining

For starters, we make a syntactic analysis of each review in the online spot review set R and obtain syntactic parsing tree. Then we begin the oword mining according to the following dependency relation: extracting verbs from dobj structure, and adjectives from nsbj and rmode. See the examples below:

E.g. 1. “我喜欢这个景区的宣传单设计。”

Dependency relationship analysis:

```
...  
asm (这个景区-3, 的-4)  
nn (设计-6, 宣传单-5)  
dobj (喜欢-2, 设计-6)
```

Feature word: “设计”

Evaluating word: “喜欢”

Structure: dobj

E.g. 2. “道路的卫生很好。”

Dependency relationship analysis:

```
assmod (卫生-3, 道路-1)
assm (道路-1, 的-2)
nsubj (很好-4, 卫生-6)
```

Feature word: “卫生”
Evaluating word: “很好”
Structure: nsubj

E.g. 3. “比较普通的一个景观”

Dependency relationship analysis:

```
advmod (普通-2, 比较-1)
rmod (景观-1, 普通-2)
...
```

Feature word: “景观”
Evaluating word: “普通”
Structure: rmod

Like the syntactic analysis mentioned above, conduct oword mining by the following steps:

- (1) Use basic transaction set of product review (r) to decide whether it contains product/service features. If it does, continue the following steps; otherwise end the whole process.
- (2) Obtain r's syntactic parsing tree.
- (3) Get the first dependency relation from the tree.
- (4) Decide if the relation identifies with dobj or nsubj dependency structure, and contains feature f or not. If yes, obtain candidate oword; otherwise skip to step (6).
- (5) If it is the dobj dependency, check expanded affective lexicon; if the nsubj dependency, check the expanded evaluating lexicon. If the result is “true”, namely the oword exists in expanded lexicon, a feature review tuple will be created; otherwise, calculate the word's similarity with every word in seed affective lexicon, and decide whether to add it to the lexicon (The similarity threshold value is set at 0.8).
- (6) If there's still dependency structure left, choose the next dependency structure and skip to step (4); otherwise, end the process.

This paper employs HowNet affective lexicon, where seed affective lexicon consists of “positive affective word (Chinese). txt” and “negative affective word (Chinese). txt”. Similarly, seed evaluating lexicon consists of “positive evaluating word (Chinese). txt” and “negative evaluating word (Chinese). txt”. Create the initial oword index table of expanded affective lexicon, words being the index, polarity being the value, negative as “-1” and positive as “1”. For the synonym dictionary designed by Harbin Institute of Technology, add synonyms in the index table to improve algorithm's recall rate.

3.2.2 Mining dword and nword

Based on the oword mining toward tourism spot's product/service, we continue to consider about dword and nword. When mining oword, extract the adverbs and negative words in the parsing tree: the former by means of advmod structure and the latter by neg structure. In the ming of adverbs of degree, use mainly “word of degree (Chinese). txt” by How Net.

E.g. 4. “景区的卫生不错，算干净，不过工作人员的态度不好了，对态度不是很满意。”

Dependency relationship analysis:

```
...
neg (满意-21, 不-18)
cop (满意-21, 是-19)
advmod (满意-21, 很-20)
...
```

To summarize, the comprehensive mining toward the order of polarity words, adverb of degree and negative words could realize “input: r—output: complete review analysis tuple”.

3.3 Lexicon's parameter setting and processing

After getting the online spot review's feature analysis, we move to the calculation of affective inclination degree. Three things should be solved one after another: the polarity parameter setting of adverbs of degree and negative words, semantic gaps caused by the different co-occurrence order of the two, and the calculation method of affective inclination.

3.3.1 Polarity parameter setting of adverbs of degree

Here we'd like to refer to the study of Xu (2007) and set different polarity parameters for 6 types of adverbs of degree on the base of HowNet degree lexicon, as shown in graph 3-1:

Table 3-1 degree adverb classification and polarity parameter list

Level	Polarity Parameter	Enumeration (N)
1.最	1.6	非常, 极其..... (N=69)
2.超	1.6	超, 过于..... (N=30)
3.很	1.4	格外, 分外..... (N=42)
4.较	1.2	更加, 较为..... (N=37)
5.稍	0.8	略微, 稍微, 挺, 蛮..... (29个)
6.欠	0.6	半点, 相对..... (12个)

3.3.2 Setting and processing of negative word parameters

For the processing of negative words, is unreasonable to simply select negative polarity, since situation will become complicated when adverbs of degree and negative words co-occur. For instance, the affective inclinations of “not very satisfied” and “very unsatisfied” are totally different.

Therefore, the author set the polarity parameter of negative words as “-0.5”, and then consider the word order of co-occurring negative words and adverbs of degree, which can be divided into two types: VDP and DNP (N-negative word, D-adverb of degree, P-polarity word).

For example, “His service attitude is very no good” is DNP structure, while “His service attitude is very not good” falls into VDP structure. The calculation method of affective inclination degree under such circumstance would be:

$$\text{deg}(DNP) = \text{deg}(d) * \text{deg}(n) * \text{deg}(p) \quad \text{Formula 3-1}$$

$$\text{deg}(NDP) = 0.8 * \text{deg}(d) * \text{deg}(p) \quad \text{Formula 3-2}$$

d stands for adverb of degree, n means negative word and p means polarity word.

4 Final affective computing and system framework integration

4.1 Final affective computing

After going through the steps above, we now define the final affective computing formula as:

(1) Specific feature tuple targeted affective computing:

$$\text{deg}(o) = \text{pol}(oword) * \text{deg}(dword) * \text{deg}(nword) \quad \text{Formula 4-1}$$

“pol (oword)” means the original polarity of polarity word; deg (dword) and deg (nword) stand for the polarity parameter of adverb of degree and negative word respectively. What is worth noticing is that when neither two exist, the result equals the default “1”; when both occur in the same context, different processing approaches should be used according to Formula 3-1 and 3-2.

(2) Specific review targeted affective computing

$$\text{deg}(r) = \frac{\sum_{i=1}^n \text{deg}(O_i)}{n} \quad \text{Formula 4-2}$$

“n” represents the number of reviews included in the tuple.

(3) Specific feature targeted affective computing

$$\text{deg}(f) = \frac{\sum_{i=j}^k \text{deg}(O_j)}{k} \quad \text{Formula 4-3}$$

“k” stands for the number of review tuples that involve a specific feature among all the online spot reviews.

(4) Tourism spot targeted affective computing:

$$\text{deg}(\text{景区}) = \frac{\sum_{i=j}^m \text{deg}(r_j)}{m} \quad \text{Formula 4-4}$$

“m” means the number of online reviews toward a specific tourism spot.

4.2 System framework integration

The affective computing system employs B/S framework with Java technology. The specific modules include:

The module of online product review collection, where WebHarvest, an open source extraction tool, is used to preset rules for review information extraction and realize automatic information collection;

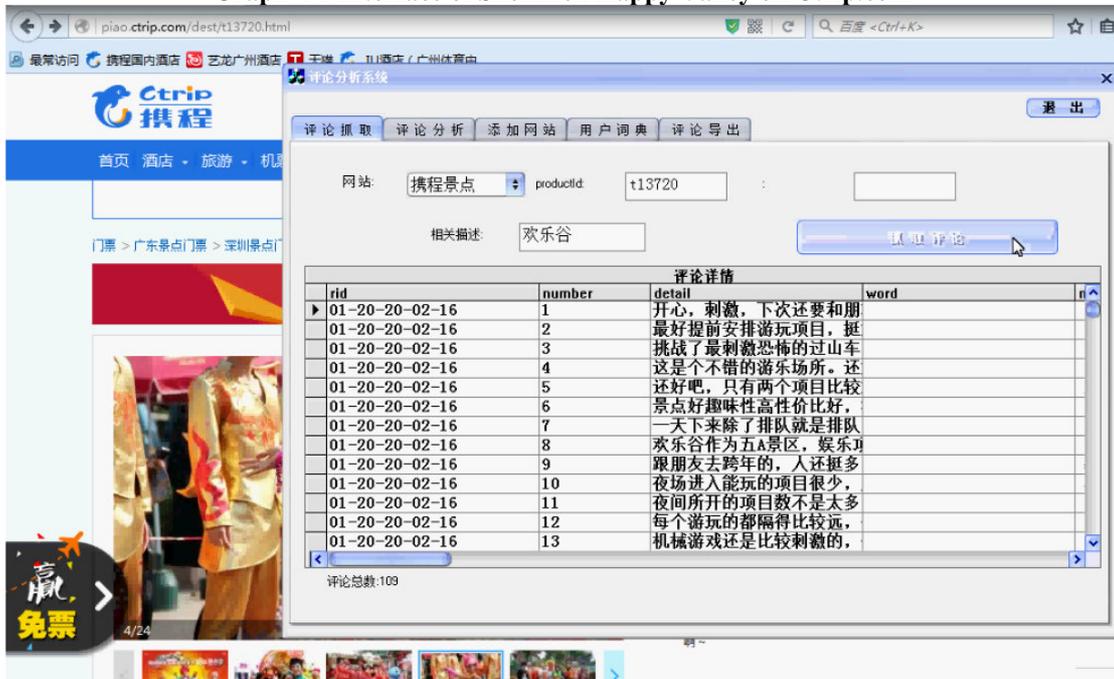
The Chinese word segmentation and part-of-speech tagging is conducted under ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System). Affective lexicon is derived from HowNet affective computing lexicon. Vocabulary similarity calculation is supported by the open source toolkit WordSimilarity from HowNet. Synonym dictionary makes reference from HIT-ITLab, and syntactic analysis is made via Stanford Parser.

We choose the interface of Shenzhen Happy Valley on Ctrip.com as the example to demonstrate the system, of which the operation procedure is shown as graphs below:

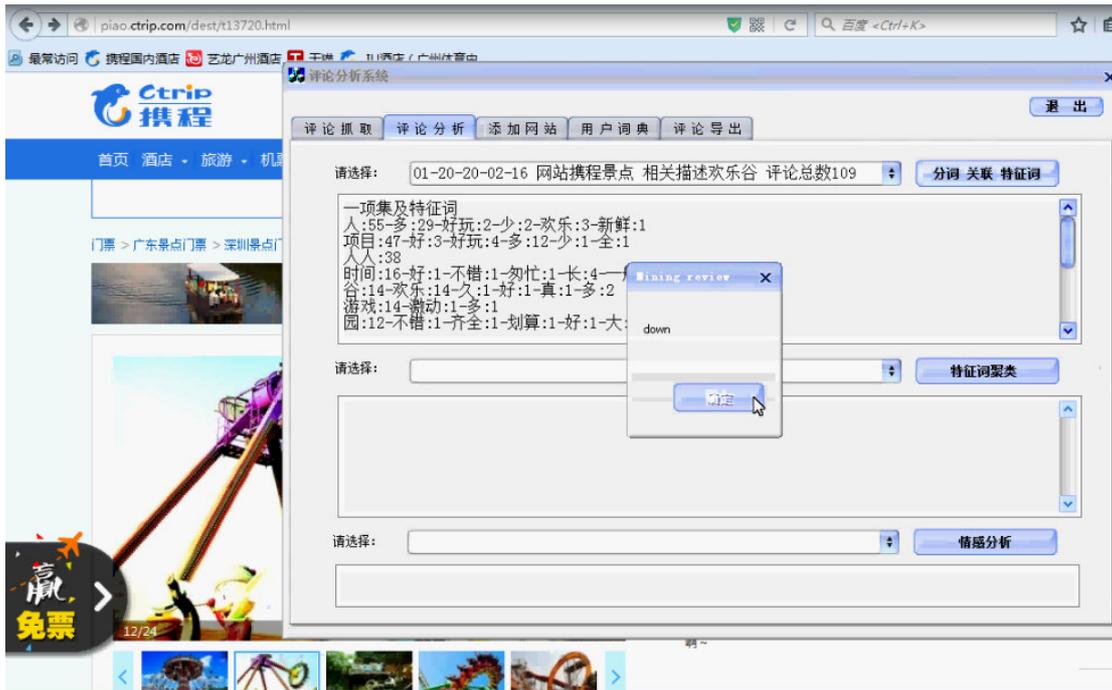
以下所有！



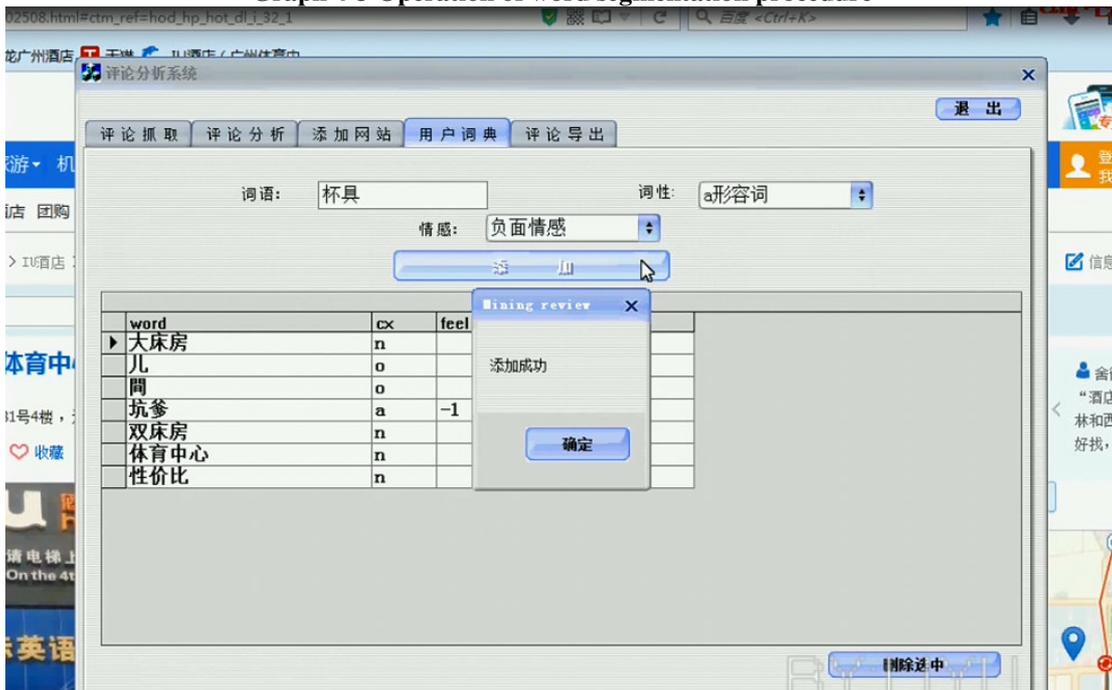
Graph 4-1 Interface of Shenzhen Happy Valley on Ctrip.com



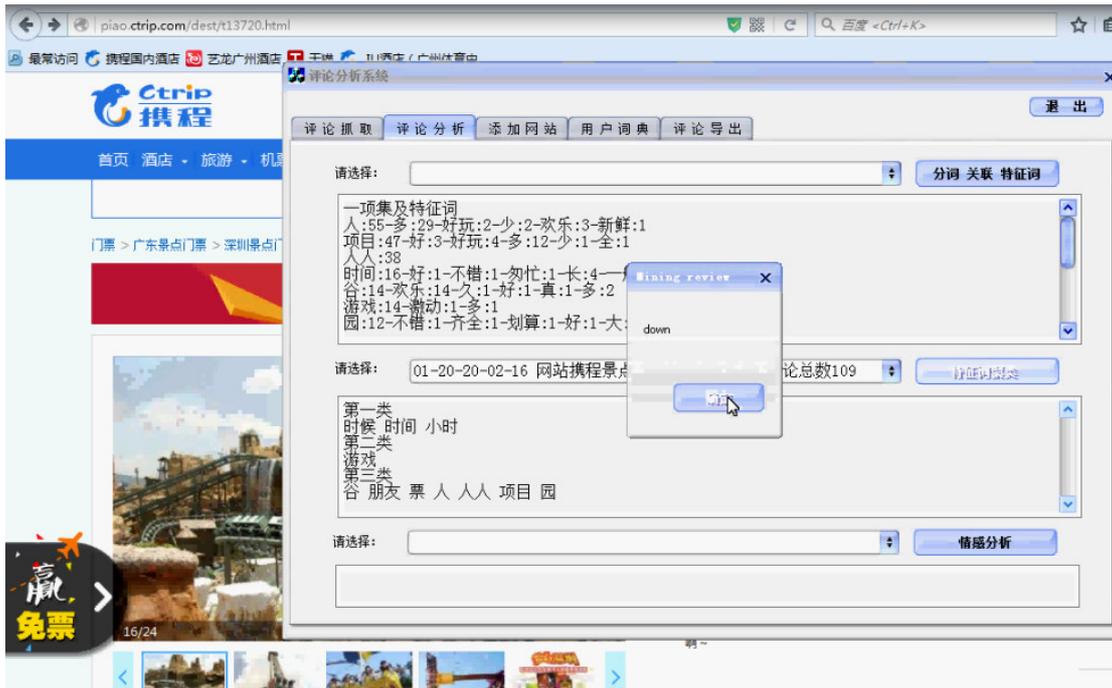
Graph 4-2 Interface of review extraction module



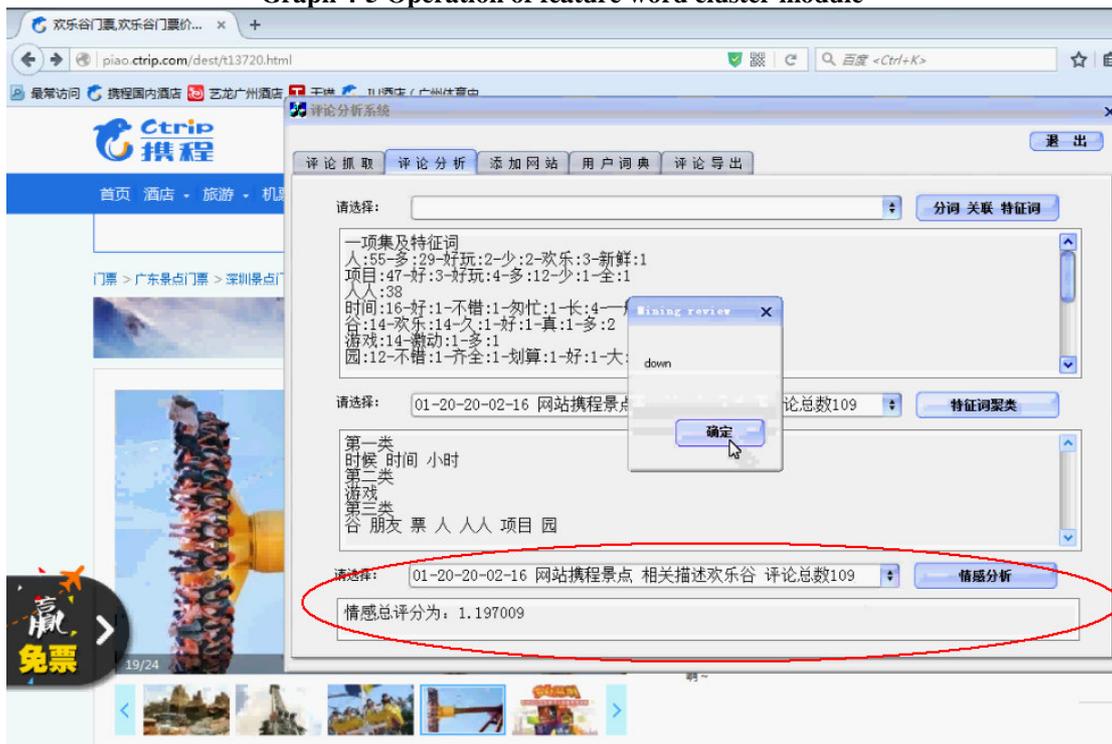
Graph 4-3 Operation of word segmentation procedure



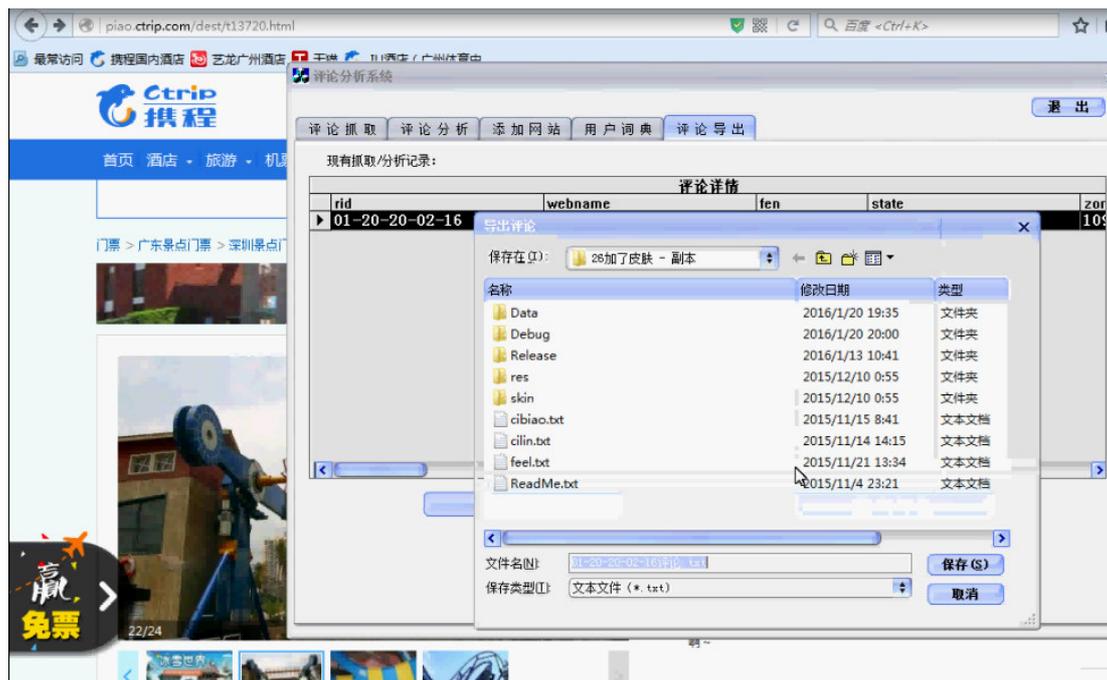
Graph 4-4 Semi-automatic update of user dictionary



Graph 4-5 Operation of feature word cluster module



Graph 4-6 Operation of Affective computing module



Graph 4-7 Operation of review export module (additional)

5 Conclusions

Based on the information redundancy of online tourism spot reviews, this paper puts forward an affective computing system for further analysis, supported by key algorithms including feature mining and affective computing. In application stage, the system can automatically collect reviews and transform complicated review information into commendatory and derogatory evaluation toward the features of tourism spots. It therefore helps tourists' travel decision-making and tourism spot managers' performance management in an effective manner.

Although the study has come up with clear operational plans of system framework and key algorithms, it still stays on theoretical designing level. In the next step, we will bring the system into practice, namely applying and examining it in real world examples, so as to further improve its accuracy and stability.

6 References

- [1] Capriello, A., Mason, P. R., Davis, B., & Crofts, J. C. (2013). Farm tourism experiences in travel reviews: A cross-comparison of three alternative methods for data analysis. *Journal of Business Research*, 66(6), 778-785.
- [2] Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, 54(3), 477-491.
- [3] Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, 19(3), 291-313.
- [4] Jeong, E., & Jang, S. S. (2011). Restaurant experiences triggering positive electronic word-of-mouth (eWOM) motivations. *International Journal of Hospitality Management*, 30(2), 356-366.
- [5] Kang, H., Yoo, S. J & Han, D. (2012). Senti-lexicon and improved Naive Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, 39(5), 6000-6010.
- [6] Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59-68.
- [7] Kim, E. E. K., Mattila, A. S., & Baloglu, S. (2011). Effects of gender and expertise on consumers' motivation to read online hotel reviews. *Cornell Hospitality Quarterly*, 52(4), 399-406.
- [8] Levy, S. E., Duan, W., & Boo, S. (2013). An analysis of one-star online reviews and responses in the Washington, D. C., lodging market. *Cornell Hospitality Quarterly*, 54(1), 49-63.
- [9] Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, 19(4), 456-474.

- [10] Lu, X., Ba, S., Huang, L., & Feng, Y. (2013). Promotional Marketing or Word-of-Mouth? Evidence from Online Restaurant Reviews. *Information Systems Research*, 24(3), 596-612.
- [11] Mauri, A. G., & Minazzi, R. (2013). Web reviews influence on expectations and purchasing intentions of hotel potential customers. *International Journal of Hospitality Management*, 34, 99-107.
- [12] 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.
- [13] Noone, B. M., McGuire, K. A., & Rohlf, K. V. (2011). Social media meets hotel revenue management: Opportunities, issues and unanswered questions. *Journal of Revenue & Pricing Management*, 10(4), 293-305.
- [14] Ogut, H., & Onur Tai, B. K. (2012). The influence of Internet customer reviews on the online sales and prices in hotel industry. *Service Industries Journal*, 32(2), 197-214.
- [15] Ogut, H., & Cezara, A. (2012). The factors affecting writing reviews in hotel websites. *Procedia-Social and Behavioral Sciences*, 58, 980-986.
- [16] Pan B., MacLaurin T., Crotts J. (2007). Travel Blogs and the Implications for Destination Marketing. *Journal of Travel Research*, 46(1):354-541.
- [17] Pantelidis, I. S. (2010). Electronic meal experience: A content analysis of online restaurant comments. *Cornell Hospitality Quarterly*, 51(4), 483-491.
- [18] Qiang Ye, Ziqiong Zhang, and Rob Law. (2009). Sentiment Classification of Online Reviews to Travel Destinations by Supervised Machine Learning Approaches. *Expert Systems with Applications*, 36(3):6527-6535.
- [19] Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32(6), 1310-1323.
- [20] Xu Linhong, Lin Hongfei, Yang Zhihao. The mechanism of text orientation recognition based on semantic understanding [J]. *Chinese Journal of information science*, 2007, 21 (1): 96-100.
- [21] Yacouel, N., & Fleischer, A. (2012). The role of cybermediaries in reputation building and price premiums in the online hotel market. *Journal of Travel Research*, 51(2), 219-226.
- [22] Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computes in Human Behavior*, 27(2), 634-639.
- [23] Ye Qiang, Zhang Ziqiong, Luo Zhenxiong. Research on the Chinese subjectivity automatic identification method for the sentiment analysis of Internet reviews. *Journal of information systems* .2007,1:79-91.
- [24] Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology & Tourism*, 10(4), 283-295.
- [25] Yoo, K. H., & Gretzel, U. (2011). Influence of personality on travel-related consumer-generated media creation. *Computes in Human Behavior*, 27(2), 609-621.
- [26] Zhang Xuefen, Li Deyu, Nan Ruixian, Wang Suge. Based on statistical methods for tourism question classification experiment research. *Development and application of computer*. 2009, 22 (1): 14-16.
- [27] Zhang, Z., Zhang, Z., Wang, F., Law, R., & Li, D. (2013). Factors influencing the effectiveness of online group buying in the restaurant industry. *International Journal of Hospitality Management*, 35, 237-245.
- [28] Zhong, L., Leung, D., Law, R., & Wu, B. (2013). eTourism in China: A review of the Chinese-language literature. *Asia Pacific Journal of Tourism Research*, 18(5), 464-482.
- [29] Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, 74(2), 133-148.