Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-018-0683-9

ORIGINAL RESEARCH

Applying sentiment analysis in social web for smart decision support marketing

Shih-Jung Wu[1] · Rui-Dong Chiang[2] · Han-Chi Chang[2]

Received: 21 August 2017 / Accepted: 7 January 2018

Abstract
Because of the rapid development of communication and service in Taiwan, competition among telecommunication companies has become ever fiercer. Differences in marketing strategy usually become the key factor in keeping existing customers while attracting new ones. Although electronic word-of-mouth (e-WOM) is one of the most important pieces of information to a consumer making a purchase decision, very few articles on opinion mining have discussed and compared the relationship between multifaceted word-of-mouth (WOM) and marketing strategy. In this paper, we use our Chinese opinion-mining system (Wu et al. in J Supercomput 73:2987–3001, 2017) not only to retrieve articles related to 4G and conduct reputation analysis but also to discuss the relation between WOM and marketing strategy. The results show that (1) e-WOM can immediately and directly reflect the results of marketing strategy, and (2) although users are primarily concerned with aspects of price, online speed, and signal quality, for most Taiwanese customers, price is the key in choosing a telecommunication company. Moreover, although this paper used 4G-related articles from June 2014 to June 2015 for analysis, the results are consistent with the Taiwanese telecommunication companies' current marketing strategy of attracting customers through low pricing.

Keywords Electronic word-of-mouth · Sentiment analysis · Opinion leaders · Online marketing

1 Introduction

Mobile communication in Taiwan has formally advanced into the 4G era. Currently, Taiwanese 4G users comprise 20% of the country's 24.63 million mobile users, and the number of telecommunication companies has increased from four to five. As 4G becomes more popular, increasing numbers of Taiwanese use handheld smart devices for a wide range of diversified services through mobile communication networks. It is foreseeable that competition among telecommunication companies will become fierce. Thus, each major telecommunication company will plan different marketing strategies to retain customers and even attempt to win other companies' customers.

With the convenience of a network, consumers can check each company's WOM through famous platforms and forums so that they can make purchasing decisions based on an understanding of each company's benefits and drawbacks with respect to network speed and quality, plan pricing, and service quality. Conversely, telecommunication companies can learn about consumers' focus through WOM related to all aspects of each telecommunication company and further understand whether other companies' marketing strategies and methods are effective in attracting consumers to extend their contracts or purchase new services. The mobile01 forum is a professional forum in Taiwan. To understand the customer's response, many marketing professionals at telecommunication companies read articles on the mobile communication discussion board on the mobile01 forum. This helps them both to understand the services that are important to customers and to propose effective marketing strategies. Therefore, in this paper, we analysed 33,080 4G-related articles on the mobile01 forum's mobile communication discussion board from June 2014 to June 2015.

✉ Shih-Jung Wu
wushihjung@mail.tku.edu.tw

1 Department of Innovative Information and Technology, Tamkang University, New Taipei, Taiwan, People's Republic of China

2 Department of Computer Science and Information Engineering, Tamkang University, New Taipei, Taiwan, People's Republic of China

Published online: 31 January 2018

Based on their granularity, opinion-mining systems can be classified into document level, sentence level, and aspect-level. Analytical results for document- and sentence-level systems are less accurate than results for aspect-level systems. However, the development and usage of document- or sentence-level systems are more convenient; therefore, Chinese opinion-mining systems for the Taiwanese market primarily involve document- or sentence-level systems. Because opinion analyses at the document and sentence levels are too coarse to determine users' opinions with precision, we have developed a semi-automatic aspect-level Chinese opinion-mining system for a specific domain (Wu et al. 2017). In this paper, we make progress in our previous work and use our opinion-mining system to analyse all 4G related articles on the mobile01 forum from June 2014 to June 2015. Next, through a comparison of the actual marketing strategy and customer e-WOM of two communication companies, we attempt to understand whether those companies' marketing strategies were designed for projects that attract customer attention.

This paper is organized into five sections. The first section introduces the research motivation and objective. The second section reviews related studies on e-WOM and opinion mining. The system architecture is introduced in the third section. The fourth section provides a detailed description of our experimental results. The fifth section contains a discussion and conclusion.

2 Related work

Opinion-mining reviews are typically analysed at various resolutions, including the following: document level, sentence level and aspect level (Liu and Zhang 2012). Document-level opinion mining identifies the overall subjectivity or sentiment about an entity as expressed in a review. Sentence-level opinion mining identifies subjective sentences and their polarities. However, as noted by (Xu et al. 2011), opinion mining at the document and sentence levels is too coarse to determine a user's opinions with precision. In this paper, we focus on aspect-level opinion mining. Aspect-level opinion mining focuses on extracting opinions about the attributes of products and services. Some related studies define opinion elements as expressing people's opinions about entities such as products, services and their attributes for the purpose of sentiment analysis (Chen et al. 2015a, b; Chen and Yao 2010; Peng and Shih 2010; Qiu et al. 2009; Wu and Wang 2014; Xu et al. 2013). There are few analogous traditional Chinese systems. L.-W. Ku, Ho, and Chen (Ku et al. 2009) have developed CopeOpi, a traditional Chinese opinion-analysis system. This system analyses articles and their opinion tendencies based on the previously established NTU sentiment dictionary. Chien-Liang, Wen-Hoar, Chia-Hoang, Gen-Chi, and Jou (Liu et al. 2012) have established a system in the field of Chinese film that allows the user to select both the name of a film to be reviewed and certain characteristics related to that film. Our previous work both established a semi-automatic aspect-level Chinese opinion-mining system that can be applied to an Internet forum and analysed the opinion trends expressed by the articles on one such forum (Wu et al. 2017). We proposed the thesaurus-based extraction of opinion elements from articles and designed algorithms for extracting and expanding the opinion elements considered by the system.

In the study of marketing, WOM is a filter; that is, information is filtered by users and transmitted amongst customers. The practicality of WOM will directly affect customers' trust-influenced behavioural intention. For many years, the theory of marketing by WOM has been studied, and WOM marketing is a marketing model evolved from WOM dissemination. WOM dissemination is an informal group influence and is defined as a verbal and non-commercial conversation about a product or brand between the disseminator and the disseminated. WOM is not quite the same as mass media: instead, it is a mouth-to-ear type of information. The disseminator of WOM is typically also a customer; through sharing purchasing and usage experience, the information is transmitted to another user, who is a future or potential customer. Moreover, because of its non-commercial character in most conditions, WOM information is typically viewed by consumers as a reliable source. In addition, some studies have discussed the relationship among brand loyalty, customer satisfaction and WOM (Huang et al. 2013; Liao et al. 2010; Takeuchi et al. 2003; Wu and Wang 2014).

With the development of the Internet era, traditional WOM marketing has been applied to the Internet. WOM is no longer purely mouth-to-ear information transferred among people. Instead, consumers can use various web platforms such as message boards, forums, BBSs, and chatrooms to disseminate their purchasing and usage experiences. In addition, consumers can obtain the desired information through web surfing, a phenomenon known as electronic word-of-mouth (e-WOM). Because the web can enable consumers to share their knowledge, experience, and opinion with high efficiency, both buyers and sellers can provide and obtain online content and engage in more active interaction. Therefore, e-WOM has four characteristics: high efficiency and low cost, high reliability, anonymity, and high interactiveness (Kaisheng 2010; Wu and Yang 2010). Unlike traditional WOM's reliance on opinion leaders, e-WOM replaces traditional WOM and can reduce consumer uncertainty and purchasing risk, serving as a communication bridge between buyers and sellers. Okada (2012) has conducted a precise analysis of hot topics in e-WOM, simulating how consumers make consumption decisions.

3 System architecture

Because opinion elements must be annotated individually for each domain (Kaisheng 2010), our system focused on a single domain for each analysis. Addressing only a single domain allows us to reduce errors by ignoring feature words or opinion words that belong to other domains. The system architecture is presented in Fig. 1. The main components of the system include the following: (1) a data collection component; (2) an opinion analysis component; and (3) a report component. The main function of the data collection component is to use a crawler to download new topics and new articles from the specified discussion area of a forum at a specified time and then to perform various pre-processing tasks, e.g., deleting portions of the article that reference articles by other authors. The opinion analysis component will analyse pre-processed articles and then

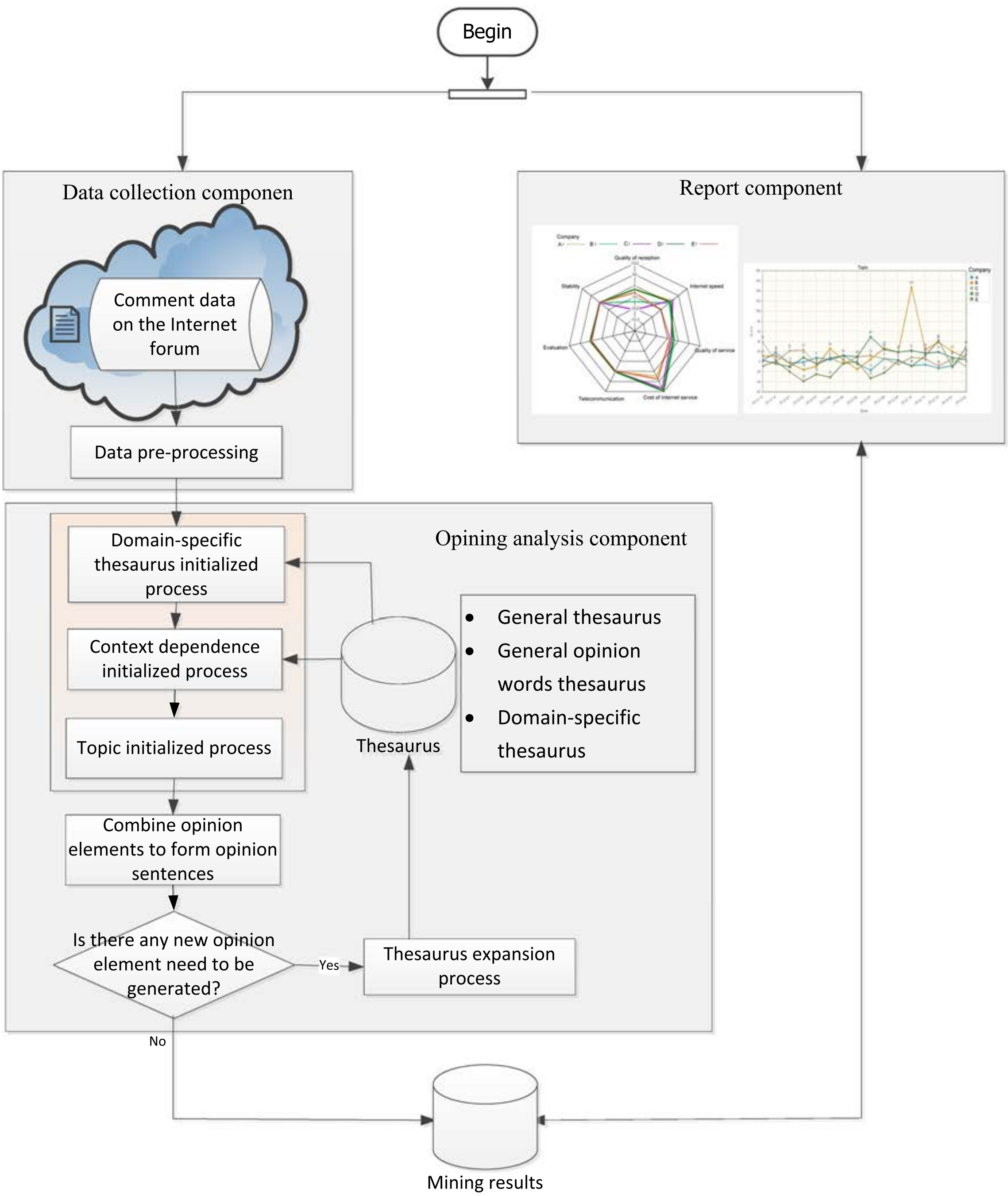

Fig. 1 System architecture

store subsequently located opinion sentences in a database. We use four types of opinion elements (company names, aspects, items, and opinion words) in the system to form an opinion sentence for the studied domain. Aspect represents a company's products or related services. If a review contains more detailed aspects, we use items to represent them. Because a sentence that expresses an opinion requires at least a topic and an opinion word to be an opinion sentence, the values of features and/or items can be null values in an opinion sentence. The report component can display results in the database through the report function.

As shown in Fig. 1, the primary components of the opinion analysis component include the following: (1) a domain-specific thesaurus initialized process, (2) a context-dependent initialized process, (3) a subject initialized process, (4) a combination of opinion elements to form opinion sentences, and (5) a thesaurus expansion process. We introduce these components briefly as follows:

Domain-specific thesaurus initialized process because there is some opinion element-related thesaurus in every field, the thesaurus used in a single field is usually limited. At this stage, some field-related opinion elements are inserted into a domain-specific thesaurus. When the opinion elements are set up, the system extracts possible opinion elements from an article based on its established general thesaurus and general opinion words thesaurus. Next, we identify through manual judgement whether the possible opinion element is a useful opinion element for a specific domain and add it to the domain-specific thesaurus.

Context-dependent process the polarity of an opinion word might be different when combined with different vocabularies. For example, the correspondence of "quality" and "high" is positive, but that of "price" and "high" is negative. In this situation, the user can set the corresponding relation between vocabularies and adjust the polarity of the correspondence.

Topic initialized process in replying, many people usually do not mention the topic of the discussion (product name). Therefore, at the time of analysis, it is difficult to know the topic discussed in the article that is receiving a reply. To resolve this issue, the system automatically analyses the topic of the article before analysing the article itself and explores the default opinion elements and potential domain-specific words from the topic. Under this topic, if the reply does not mention the topic of discussion, the pre-set product name of the topic will be used as the topic of the reply. If a new opinion vocabulary appears in the topic, the user can add the new vocabulary into the thesaurus and implement the relevant settings.

Combination of opinion elements to form opinion sentences because a language's syntax rules are relatively basic and static (Liu and Zhang 2012; Xu et al. 2011), in this paper, the opinion elements of articles were extracted based on lexicons and combined with the sentence patterns (general sentences, equivalent sentences, and comparative sentences) and nearby approach to form opinion sentences.

Thesaurus expansion process two types of cases may generate new opinion elements. First, new Internet words periodically emerge on Internet forums. When telecommunication companies introduce new products, these new words and product names may become new opinion elements. Therefore, when any opinion element's value is a null value, the rule will trigger the word-hyphenation algorithm to extract new words that may become new opinion elements. Second, in Chinese, when an opinion word (referred to as *OP*) is combined with a specific word or concatenated with other opinion words, it may generate a change in the polarity or meaning of the expressed opinion. Consequently, to process this change, we use some rules to address these situations.

4 E-WOM of telecommunication companies and their marketing decision analyses

In this paper, we will perform e-WOM analysis of 33,880 4G-related articles on the mobile communication discussion board of the mobile01 forum from June 2014 to June 2015. We will explore the relationship between changes in consumers' e-WOM and marketing strategy through the Taiwan 4G market's e-WOM on the mobile01 forum. To avoid being influenced by multiple articles by the same person, every person only counts once for the same company name, aspect name or item name within the same time period. The grading algorithm is shown in Fig. 2.

4.1 Comprehensive evaluation and analysis of telecommunication companies

Because Company E, Company D, and Company C changed their marketing strategy in March 2015, it is easy to see these companies' change in web e-WOM after the strategy change. We divide the region of e-WOM into two segments: the first segment is from June 2014 to February 2015 (Fig. 3), and the second segment is from March 2015 to June 2015 (Fig. 4). Figures 3 and 4 are the comprehensive analysis diagram of telecommunication companies in the segments of June 2014 to February 2015 and March 2015 to June 2015, respectively:

Taiwan currently has five telecommunication companies. Company E is the largest and has the most customers. Company B practised a low-price marketing strategy after it launched in August 2014; as shown in Figs. 3 and 4, the positive and negative assessment numbers for these two companies are higher than those for the other three

Fig. 2 The grading algorithm

Grading algorithm

Input: Complete sentences / Analysis targets
Analysis targets= Topic(T)/Aspect(A)/Item(I)
Output: Grading result of analysed targets
(Note: if the same complete sentences from the same post article, **only counted once**)

switch(Analysis targets)
case T:
Set S: Complete sentence belongs to specific form as follow: Topic *T* + any Aspect/ ɸ + any Item/ ɸ + OP
case A:
Set S: Complete sentence belongs to specific form as follow: Topic *T* + Aspect *A* + any Item/ ɸ + OP
case I:
Set S: Complete sentence belongs to specific form as follow: Topic T + Aspect A + Item I + OP

$$\mathrm{P}_S^i = \begin{cases} +1, & \text{Positive opinion sentence for reviewer } i \text{ (complete sentence has positive OP)} \\ -1, & \text{Negative opinion sentence for reviewer } i \text{ (complete sentence has negative OP)} \end{cases}$$

for analysed target.
Reviewer$_i$: represents the opinion tendency of all reviewer *i*'s posts for analysed target
G: represents the final grades of analysed target.

$$\mathrm{Reviewer}_i = \begin{cases} +1, & if \sum P_S^i > 0 \\ 0, & if \sum P_S^i = 0 \\ -1, & if \sum P_S^i < 0 \end{cases}$$

Return G=$\sum Reviewer_i$

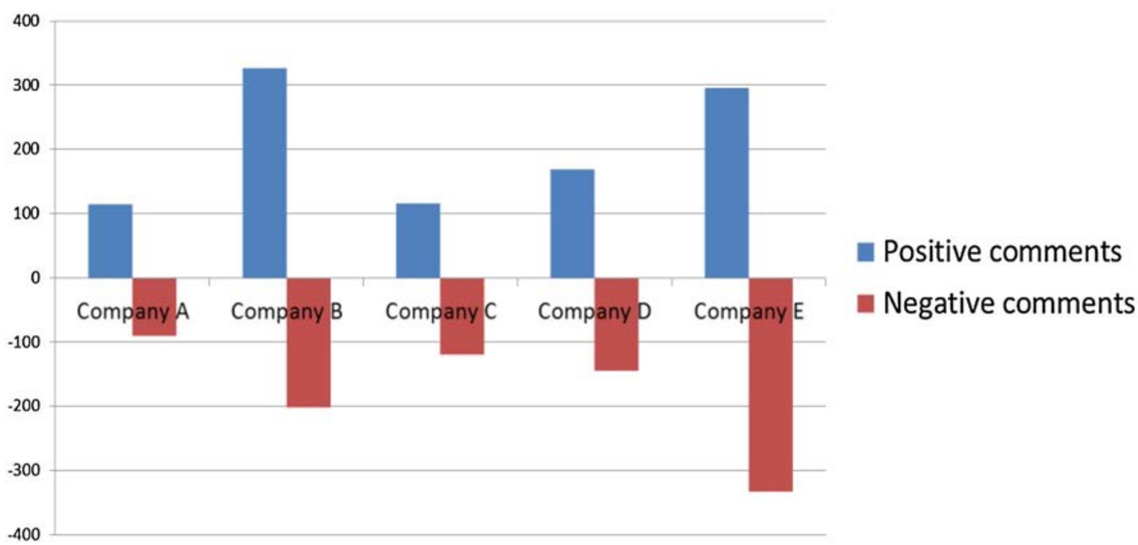

Fig. 3 Comprehensive analysis diagram of telecommunication companies (2014/06–2015/02)

companies. Moreover, we also obtain the following findings. First, significantly more people make positive assessments about Company E than negative ones starting in March 2015. In comparison, there were more negative assessments than positive ones before March of 2015. Similarly, it is obvious that the assessments of Company C and Company D improved after March 2015. s, after it began formal operations, Company B always received many more positive assessments then negative ones. Because the reason for Company C, Company D and Company E's improved assessments after March 2015 is the same and Company E is the largest telecommunication company with the earliest 4G marketing commercials and the most customers, we will compare Company E with Company B to discuss the reason for the improved assessments of Company E starting in March 2015.

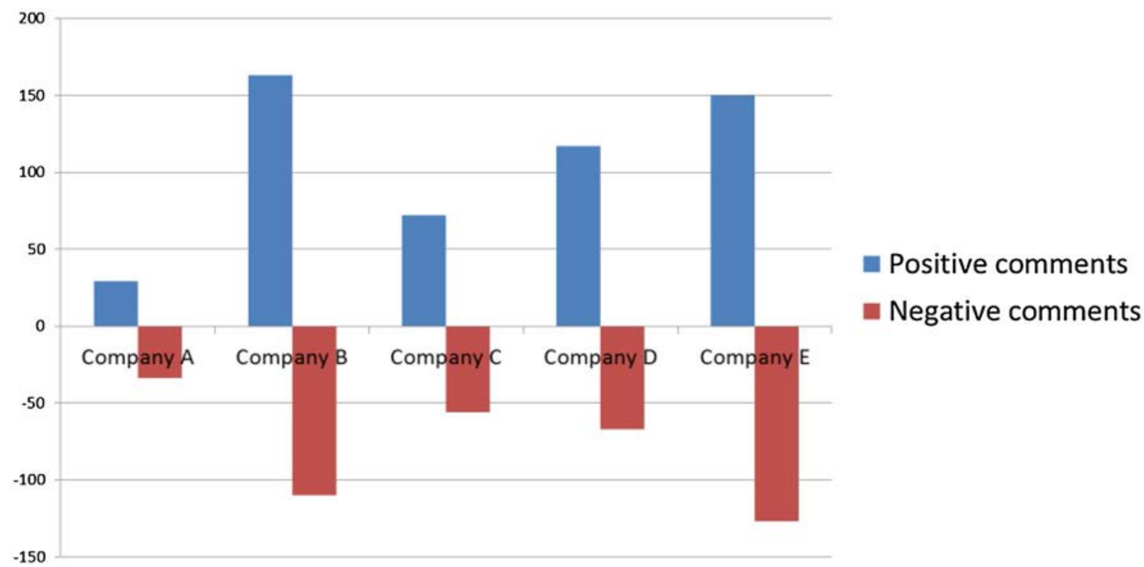

Fig. 4 Comprehensive assessment of telecommunication companies (2015/03–2015/06)

4.2 Marketing strategy and assessment analysis

This section conducts an analysis and comparison of trends in the overall assessment of the current situations of Company B and Company E and explores the reason for the data shown in Fig. 4. Figure 5 is the trend in the number of discussion participants with respect to each aspect. Figure 5 shows that online speed has been discussed by the largest number of people. Originally, signal quality was the topic that attracted the second-highest number of people, but it was overtaken by pricing after March 2015. Figure 5 shows that users primarily focus on the three aspects of price, online speed, and signal quality. However, signal quality is not related to Company B and Company E's marketing strategies. Therefore, we only consider the relationship between e-WOM and the two aspects of price and online speed in the two companies' marketing strategies.

Figures 6 and 7 show the price, online speed, and signal quality e-WOM percentage distribution of Company B and Company E in the two different segments. From June 2014 to February 2015, Company E's marketing strategy emphasized 4G online speed to attract customers, whereas Company B used a low-price strategy to attract customers. As shown in Fig. 6, the positive assessment ratio for Company

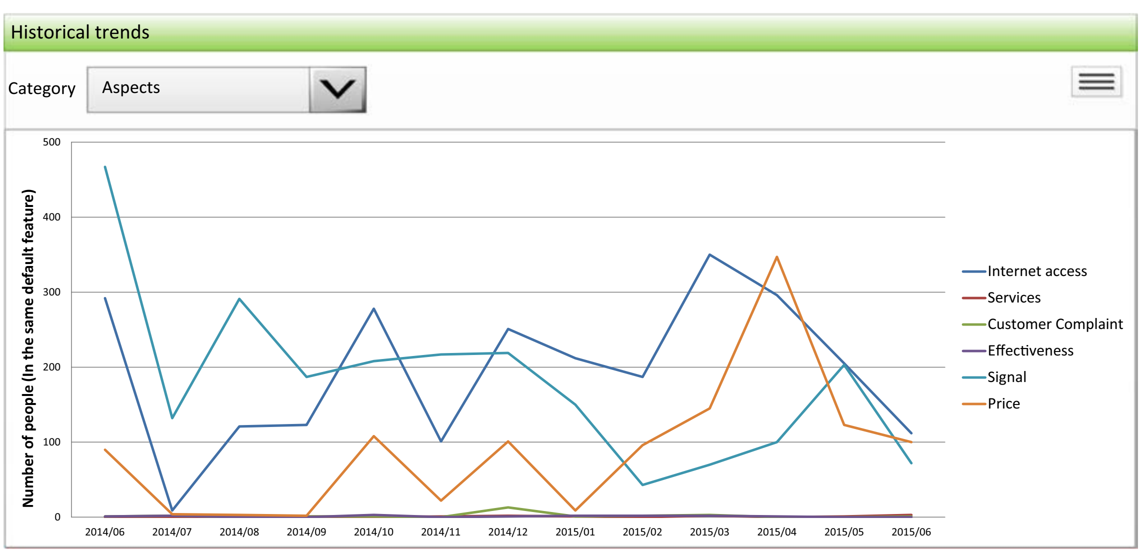

Fig. 5 Trend of the number of discussion participants in each aspect

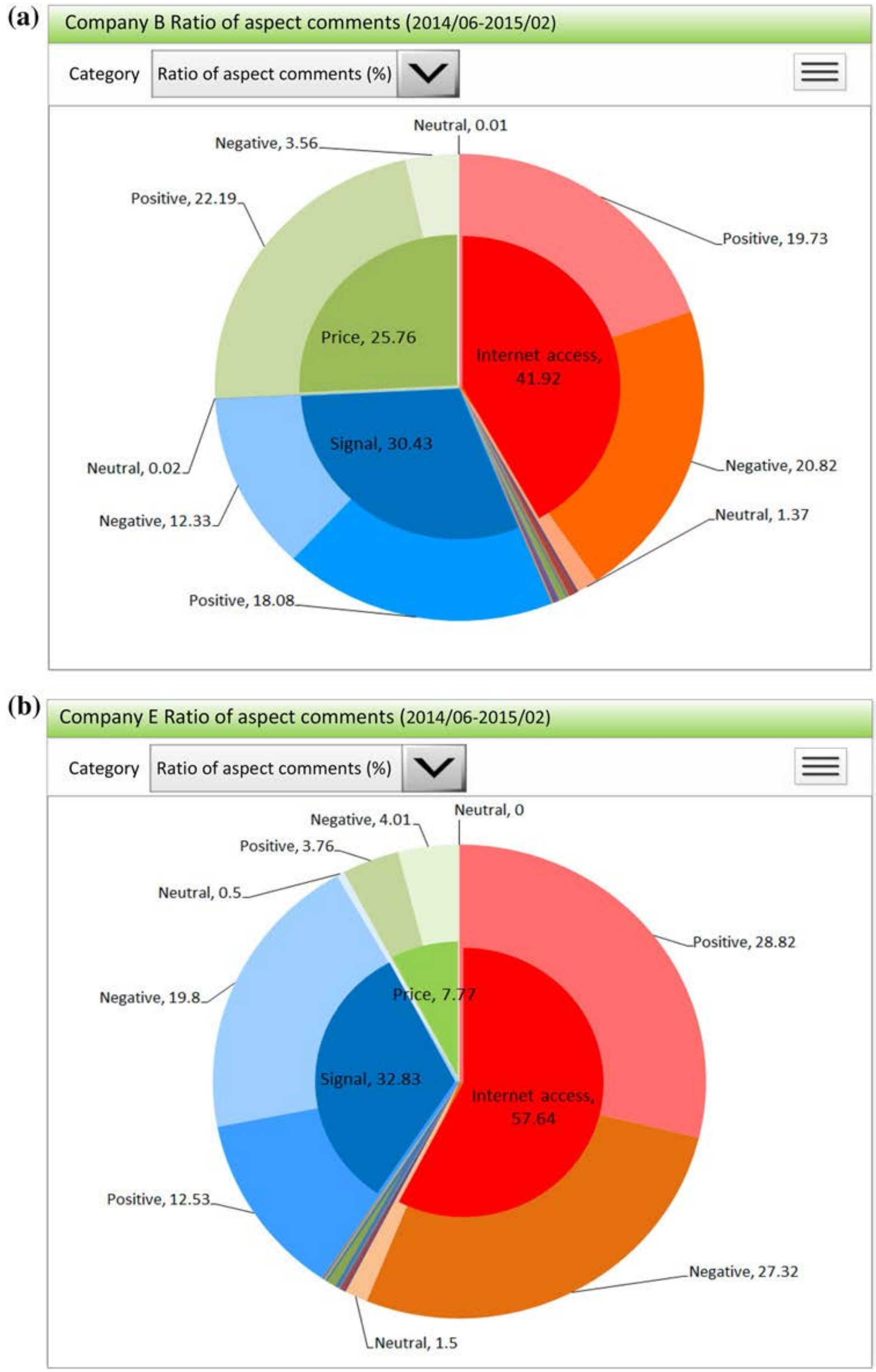

Fig. 6 **a** The distribution of assessments for Company B (201,406–201,502). **b** The distribution of assessments for Company E (201,406–201,502)

E's online speed is 28.82%, whereas the negative assessment ratio is 27.32%, slightly lower than the positive assessment ratio. However, although the negative assessment ratio for Company B's online speed is higher than the positive ratio, the difference is very small, 19.73 and 20.82%. The distribution results showed that there is no significant difference in the two companies' e-WOM in terms of online speed. This result shows that Company E's initial focus on online speed was not very successful. However, the positive assessment of Company B in terms of price is the highest at 22.19%, whereas the negative assessment is only 3.56%. Conversely, the positive and negative assessments of Company E's price

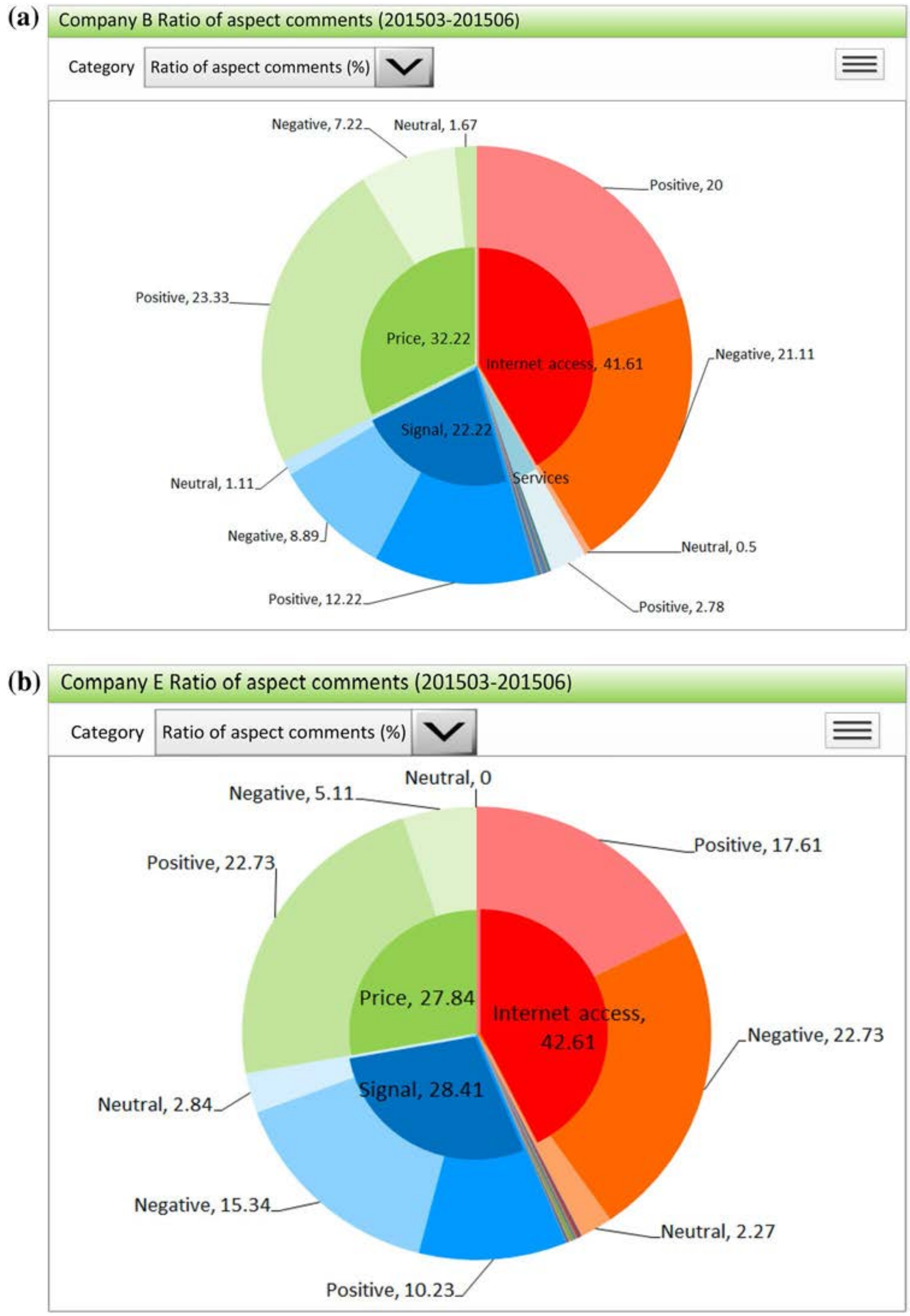

Fig. 7 **a** The distribution of assessments for Company B (201,503–201,506). **b** The distribution of assessments for Company E (201,503–201,506)

are 4.01 and 3.76% respectively; this result indicates that Company B's pricing strategy was successful.

Starting in March 2015, Company E changed its strategy of attracting customers by emphasizing 4G online speed, introducing a new low pricing programme. At the same time, Company C and Company D followed suit and introduced new low pricing programmes. Company B scrapped its original price increase plan and maintained its low price level. As shown in Fig. 7, when Company E switched to the low-price marketing strategy, although positive e-WOM on online speed decreased slightly, the positive assessment of price increased dramatically from 4.01% in June 2014 to 22.73% in February 2015, whereas the negative assessment of price increased only slightly from 3.76 to 5.11%. Even though Company B's price remains 60% lower than that of other companies, because its price remained the same, the positive assessment of Company B in terms of pricing increased only slightly from 22.19 to 23.33%, whereas the negative assessment doubled from 3.56

to 7.22%. After Company E changed its pricing strategy, pricing e-WOM very obviously showed that Company E successfully used pricing to attract customers' attention and positive assessment improved dramatically. Thus, we can see that Company E's price-changing strategy was successful. Meanwhile, through the monthly pricing e-WOM diagram similar to Fig. 3, we can see that the number of people positively assessing Company E's price immediately surpassed the number of people positively assessing Company B's in April 2015. From the above analysis, we can see that from a consumer standpoint, pricing can indeed attract consumers' attention. Indeed, as of December 2016, Company C, Company D, and Company E are still practising the marketing strategy of attracting customers with low prices. The current price level is only at 50% of promotional prices in March 2015. Thus, it can be seen that the effect of the marketing strategy will immediately be reflected in the e-WOM performance.

5 Conclusion and discussion

We have performed analysis on e-WOM for service products in 4G markets. The results indicated that price can indeed attract consumers' attention. In December 2016, Company C, Company D, and Company E are still practising the marketing strategy of attracting customers with low prices; the results also showed that the effects of marketing strategy will immediately be reflected in e-WOM performance. Because a user's online opinions (WOM) about a company's product or service reflect his or her true feelings, companies can use e-WOM to examine the effectiveness of their marketing strategy and make corresponding adjustments. Moreover, they can also identify deficiencies of their own products and services and make improvements.

Because only Company E has the mobile communication range to cover all of Taiwan, customers naturally think that Company E's signal quality is reliable. However, we can see from Figs. 6 and 7 that the performance of Company E's signal quality is not ideal. After analysis, we discovered that many users complained about Company E's indoor reception quality; many people gave negative assessments after reporting relevant situations such as the lack of an indoor signal. This is not inconsistent with the consensus of most consumers because the consensus is that Company E's outdoor signal is reliable. Upon discovering this issue through e-WOM, Company E has recently invested substantial resources to improve its indoor reception.

References

Chen M, Yao T (2010) Combining dependency parsing with shallow semantic analysis for Chinese opinion-element relation identification. In: 2010 4th International Universal Communication Symposium, 18–19 Oct. 2010. pp 299–305. https://doi.org/10.1109/IUCS.2010.5666009

Chen J, Huang DP, Hu S, Liu Y, Cai Y, Min H (2015a) An opinion mining framework for Cantonese reviews. J Ambient Intell Humaniz Comput 6:541–547. https://doi.org/10.1007/s12652-014-0237-8

Chen Y-W, Wang J-L, Cai Y-Q, Du J-X (2015b) A method for Chinese text classification based on apparent semantics and latent aspects. J Ambient Intell Humaniz Comput 6:473–480. https://doi.org/10.1007/s12652-015-0257-z

Huang L, Shao J, Wang W (2013) Research on the relationships between hotel internet word-of-mouth and customers' behavior intention based on trust. In: 2013 6th International conference on information management, innovation management and industrial engineering, 23–24 Nov. 2013. pp 250–254. https://doi.org/10.1109/ICIII.2013.6703132

Kaisheng ZO (2010) Buying impulsiveness and word-of-mouth online in the e-marketing age: a theoretical model approach. In: 2010 International conference on management of e-Commerce and e-Government, 23–24 Oct. 2010. pp 133–136. https://doi.org/10.1109/ICMeCG.2010.34

Ku L-W, Ho H-W, Chen H-H (2009) Opinion mining and relationship discovery using CopeOpi opinion analysis system. J Am Soc Inform Sci Technol 60:1486–1503. https://doi.org/10.1002/asi.21067

Liao SH, Chung YC, Hung YR, Widowati R (2010) The impacts of brand trust, customer satisfaction, and brand loyalty on word-of-mouth. In: 2010 IEEE international conference on industrial engineering and engineering management, 7–10 Dec. 2010, pp 1319–1323. https://doi.org/10.1109/IEEM.2010.5674402

Liu B, Zhang L (2012) A survey of opinion mining and sentiment analysis. In: Aggarwal CC, Zhai C (eds) Mining text data. Springer US, Boston, pp 415–463. https://doi.org/10.1007/978-1-4614-3223-4_13

Liu CL, Hsaio WH, Lee CH, Lu GC, Jou E (2012) Movie rating and review summarization in mobile environment. IEEE Trans Syst Man Cybern Part C (Applications Reviews) 42:397–407. https://doi.org/10.1109/TSMCC.2011.2136334

Okada I (2012) Relationship between online word-of-mouth communication and consumer behavior. In: 2012 45th Hawaii international conference on system sciences, 4–7 Jan. 2012, pp 1807–1813. https://doi.org/10.1109/HICSS.2012.500

Peng TC, Shih CC (2010) Using Chinese part-of-speech patterns for sentiment phrase identification and opinion extraction in user generated reviews. In: 2010 fifth international conference on digital information management (ICDIM), 5–8 July 2010, pp 120–127. https://doi.org/10.1109/ICDIM.2010.5664631

Qiu G, Liu B, Bu J, Chen C (2009) Expanding domain sentiment lexicon through double propagation. In: Paper presented at the Proceedings of the 21st international jont conference on Artifical intelligence, Pasadena, California, USA

Takeuchi S, Kamahara J, Shimojo S, Miyahara H (2003) Human-network-based filtering: the information propagation model based on word-of-mouth communication. In: 2003 Symposium on applications and the internet, 2003. Proceedings, 27–31 Jan. 2003, pp 40–47. https://doi.org/10.1109/SAINT.2003.1183031

Wu X, Wang J (2014) Micro-blog in China: identify influential users and automatically classify posts on Sina micro-blog. J Ambient Intell Humaniz Comput 5:51–63. https://doi.org/10.1007/s12652-012-0121-3

Wu T, Yang X (2010) Electronic word of mouth in online social networks. In: 2010 Second International Conference on Communication Systems, Networks and Applications, June 29 2010–July 1 2010, pp 249–253. https://doi.org/10.1109/ICCSNA.2010.5588871

Wu S-J, Chiang R-D, Ji Z-H (2017) Development of a Chinese opinion-mining system for application to Internet online forums. J

Supercomput 73:2987–3001. https://doi.org/10.1007/s11227-016-1816-6
Xu K, Liao SS, Li J, Song Y (2011) Mining comparative opinions from customer reviews for competitive intelligence. Decis Support Syst 50:743–754. https://doi.org/10.1016/j.dss.2010.08.021
Xu G, Huang C-R, Wang H (2013) Extracting Chinese product features: representing a sequence by a set of skip-bigrams. In: Ji D, Xiao G (eds) Chinese lexical semantics: 13th workshop, CLSW 2012, Wuhan, China, July 6–8, 2012, revised selected papers. Springer, Berlin, pp 72–83. https://doi.org/10.1007/978-3-642-36337-5_9