Crowd Intelligence: Conducting Asymmetric Impact-Performance Analysis based on Online Reviews

Jian-Wu Bi
Nankai University

Yang Liu
Northeastern University

Zhi-Ping Fan
Northeastern University

***Abstract*—Asymmetric impact-performance analysis (AIPA) is an effective technique for understanding customer satisfaction (CS) and formulating improvement strategies for products and services. Typically, AIPA is conducted based on data obtained from customer surveys, which are expensive in terms of time and money. As a new data source, online reviews have many advantages, which are a promising data source for conducting AIPA. To this end, this paper proposes a method for conducting AIPA based on online reviews. To illustrate the feasibility and validity of the proposed method, a case study of AIPA for a five-star hotel in Singapore is given. The proposed method can provide managers one more choice for conducting AIPA with lower cost and shorter time since products and services online reviews can be easily collected.**

UNDERSTANDING customer satisfaction (CS) and formulating improvement strategies for products and services are crucial to the success of a company. A popular tool to do so is importance-performance analysis (IPA), which maps product and service attributes on a two dimensional grid according to the attributes performance and importance [1]. IPA assumes an attribute's positive performance and its negative performance have the same impact on CS when they have the same size. In other words, the relationship between an attribute's performance and CS is assumed to be symmetrical in IPA. However, the assumption has been criticized and proven not always holding true [2], [3]. In addition, to understand and analyze CS more effectively, attributes should be classified into different categories, such as basic attributes, excitement attributes, and performance attributes [4]. To this end, AIPA is proposed based on the three-factor theory of CS [2]. By employing AIPA, both the asymmetric influences of attributes on CS and their performances can be

  This article is part of the IEEE IS ACSA series

considered at the same time [2], [4].

Typically, AIPA is conducted based on the data obtained from customer surveys, which require well-designed questionnaires and reasonable implementation processes to ensure the quality of the data [5]. Obviously, such data collection approach is expensive in terms of time and money. Therefore, it is worthwhile to consider other alternative sources of data for conducting AIPA. In recent years, massive online reviews have emerged on the Internet, which contain a lot of valuable information (e.g., customers concerns and opinions) [6], [7]. The information embedded in these online reviews is valuable for understanding CS, mining customer preference and formulating products and services improvement strategies [4], [8]. Compared with the data obtained from customer surveys, online reviews are publicly available and can be easily collected. Therefore, online reviews are undoubtedly an emerging and promising source of data for conducting AIPA. However, studies on conducting AIPA based on online reviews have not been found.

The objective of this paper is to propose a method for conducting AIPA based on online reviews. In the method, customers' sentiments toward attributes are first mined from online reviews. Then, in accordance with the obtained customer sentiments, the effects of customers' sentiments toward attributes on CS are measured. On this basis, the asymmetric impact of each attribute on CS is estimated, and the performance of each attribute is calculated. Based on the obtained asymmetric impact and performance of each attribute, AIPA can be conducted, and products and services improvement strategies can be formulated. Finally, a case study of a five-star hotel is given to illustrate the feasibility and validity of the proposed method.

A BRIEF INTRODUCTION TO AIPA

AIPA is an effective technique for understanding CS and formulating improvement strategies for products and services. An example of the AIPA plot is given in Figure 1. It can be seen from Figure 1 that in accordance with the attribute's performance, the attributes are classified into two categories, i.e., high performance and low performance. An attribute belongs to high performance, indicating that the performance of the attribute is higher than the average performance of all attributes. Similarly, an attribute belongs to low performance, indicating that the performance of the attribute is lower than the average performance of all attributes. Meanwhile, according to the attributes impact asymmetry index, the attributes can be also classified into three categories, i.e., performance attributes, excitement attributes and basic attributes, as shown in Figure 2. Meanings of the three categories of attributes can be found in Bi et al. [4] Therefore, according to each attributes performance and each attributes impact asymmetry index, attributes of products and services are classified into six categories, i.e., Low performance-Excitement attributes (LE), Low performance-Performance attributes (LP), Low performance-Basic attributes (LB), High performance-Excitement attributes (HE), High performance-Performance attributes (HP) and High performance-Basic attributes (HB). When formulating improvement strategies for products and services considering limited resources of the company, the priority order of these six categories of attributes is: LB $>$ LP $>$ LE $>$ HB $>$ HP $>$ HE.

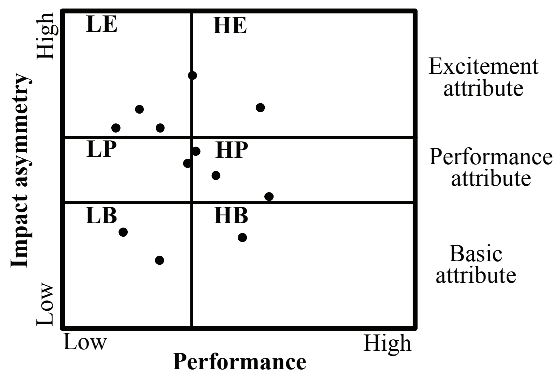

Figure 1. An example of the AIPA plot.

METHOD

In this section, a method for conducting AIPA based on online reviews is given. The framework of the proposed method is shown in Figure 3, which consists of three stages, i.e., (1) mining customers' sentiments toward attributes from online reviews, (2) measuring the effects

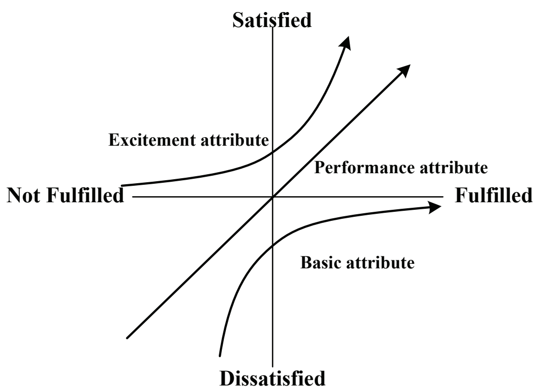

Figure 2. Classification of attributes in the three-factor theory.

of customer sentiments toward attributes on CS, and (3) conducting AIPA. Detailed descriptions of the three stages will be respectively given in the subsequent subsections.

Mining customers' sentiments toward attributes from online reviews

This stage mainly includes two steps, i.e., (1) Extracting the important attributes of products or services from online reviews, and (2) Identifying customers sentiments toward the important attributes. For the first step, following Bi et al. [4], the latent dirichlet allocation (LDA) topic model is adopted to extract the important attributes. For the second step, the SenticNet 5 proposed by Cambria et al. [9] is employed for identifying customers sentiments (polarities and intensities) toward the extracted attributes. It should be noted that other attributes extraction methods and sentiment analysis methods can also be adopted to implement the above process [10], [11].

Measuring the effects of customers' sentiments toward attributes on CS

For measuring the effects of customers' sentiments toward attributes on CS, the ensemble neural network based model (ENNM) proposed by Bi et al. [4] is employed. According to the obtained positive and negative sentiment polarities, two kinds of effects can be obtained using the ENNM, i.e., the effect of customer's positive sentiment toward attribute A_i on CS (denoted as $\bar{E}_i^{\mathrm{Pos}}$) and the effect of customer's negative sentiment toward attribute A_i on CS (denoted as $\bar{E}_i^{\mathrm{Neg}}$), $i = 1, 2, ..., I$.

Conducting AIPA

To conduct AIPA, the performance of each attribute and the asymmetric impact of each attribute on CS need to be estimated. Then, according to the obtained attribute's performance and asymmetric impact, the AIPA plot can be constructed. The details are given below.

Estimating the performance of each attribute Following Bi et al. [8], the performance of A_i can be estimated by calculating the average value of the obtained customers sentiment intensities related to A_i, denoted as Per_i, $i = 1, 2, ..., I$.

Estimating the asymmetric impact of each attribute on CS According to the definitions of $\bar{E}_i^{\mathrm{Pos}}$ and $\bar{E}_i^{\mathrm{Neg}}$, we know $\bar{E}_i^{\mathrm{Pos}}$ can be regarded as the effect of A_i on CS when customer's requirement on A_i is fulfilled; conversely, $\bar{E}_i^{\mathrm{Neg}}$ can be regarded as the effect of A_i on CS when customer's requirement on A_i is not fulfilled, $i = 1, 2, ..., I$.

Let Range_i denote an attribute's range of impact on CS, $i = 1, 2, ..., I$. Range_i can be calculated by Equation (1).

$$\mathrm{Range}_i = \left|\bar{E}_i^{\mathrm{Pos}}\right| + \left|\bar{E}_i^{\mathrm{Neg}}\right|, i = 1, 2, ...I. \quad (1)$$

Let AI_i denote the asymmetric impact (AI) of attribute A_i on CS, $i = 1, 2, ..., I$. According to Caber et al. [2], AI_i can be calculated by Equation (2), i.e.,

$$\mathrm{AI}_i = \frac{\left|\bar{E}_i^{\mathrm{Pos}}\right| - \left|\bar{E}_i^{\mathrm{Neg}}\right|}{\mathrm{Range}_i}, i = 1, 2, ...I. \quad (2)$$

Obviously, $-1 \leq \mathrm{AI}_i \leq 1$, $i = 1, 2, ...I$.

(i) If $\mathrm{AI}_i = -1$ (i.e., $\bar{E}_i^{\mathrm{Pos}} = 0$), then it means CS will decrease when customer's requirement on A_i is not fulfilled, but CS will not increase when customer's requirement on A_i is fulfilled.

(ii) If $\mathrm{AI}_i = 0$ (i.e., $\left|\bar{E}_i^{\mathrm{Pos}}\right| = \left|\bar{E}_i^{\mathrm{Neg}}\right|$), then it means that the increase degree and the decrease degree of CS are equally when customer's requirement on A_i is fulfilled or not fulfilled.

(iii) If $\mathrm{AI}_i = 1$ (i.e., $\bar{E}_i^{\mathrm{Neg}} = 0$), then it means CS will increase when customer's requirement on

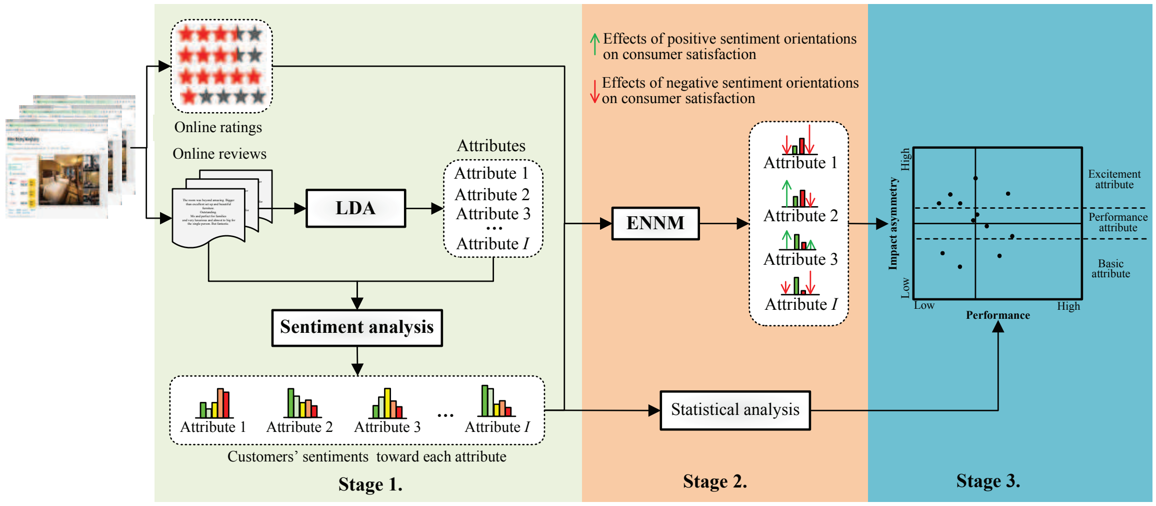

Figure 3. The framework for conducting AIPA based on online reviews.

A_i is fulfilled, but CS will not decrease when customer's requirement on A_i is not fulfilled.

To classify the attributes into different categories, a cut-off point θ is defined subjectively, $0 < \theta < 1$. Then, the category of each attribute can be determined. Specifically, if $-1 \leq A_i < -\theta$, then A_i is regarded as a basic attribute; if $-\theta \leq A_i \leq \theta$, then A_i is regarded as a performance attribute; if $\theta < A_i \leq 1$, then A_i is regarded as an excitement attribute, $i = 1, 2, ...I$.

Constructing the AIPA plot In accordance with the obtained Per_i and AI_i, the AIPA plot can be drawn with AI_i on the vertical axis and Per_i on the horizontal axis, $i = 1, 2, ...I$. The AIPA plot can help managers to understand CS and formulate improvement strategies for the product or service by allocating the limited resources.

CASE STUDY

To illustrate the feasibility and validity of the proposed method, a case study is given in this section.

Data

We take a five-star hotel in Singapore as the research object. Related data of the hotel are collected from Tripadvisor (https://www.tripadvisor.com), one of the world's leading tourism websites. A total number of 20,134 online reviews were collected. After removing invalid and non-English reviews, we obtained 14,652 valid online reviews. The collected reviews include 2,252,172 words, and the average number of words in each review is about 155.

Mining customers' sentiments toward attributes from online reviews

According to the process for extracting important attributes, we obtain nine attributes, i.e., value (A_1), location and transport (A_2), cleanliness (A_3), service (A_4), room (A_5), food and drink (A_6), check in and out (A_7), facility (A_8), wifi and internet (A_9). Using the SenticNet 5 proposed by Cambria et al. [9], the sentiment intensity of each sentence concerning each attribute is determined. The results are shown in Table 1.

Measuring the effects of customers' sentiments toward attributes on CS

Using the ENNM proposed by Bi et al. [4], the values of $\bar{E}_i^{\mathrm{Pos}}$ and $\bar{E}_i^{\mathrm{Neg}}$ can be obtained, as shown in Table 2.

Conducting AIPA

By calculating the average value of customers sentiment intensities related to attribute A_i , each attributes performance of the five-star hotel can be obtained, as shown in Table 3.

According to the obtained $\bar{E}_i^{\mathrm{Pos}}$ and $\bar{E}_i^{\mathrm{Neg}}$, Range_i and AI_i can be respectively calculated by Equations (1) and (2), the results are shown in Table 4.

Following Caber et al. [2], the value of θ is set to 0.1 in this study. In accordance with the

Table 1. The sentiment intensity of each sentence concerning each attribute.

Review	Attribute									CS
	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	
r_1	-0.51	0	-0.912	0	-0.357	0	0	0	0	10
r_2	0	0.887	0.137	0	0	0.857	0.329	0	0	50
...	...	...	...	...	...	...	...	...	...	...
$r_{14,652}$	0	0.091	-0.05	0	0	0.357	0	0	0.076	30

Table 2. The values of $\bar{E}_i^{\text{Pos}}$ and $\bar{E}_i^{\text{Neg}}$.

	$\bar{E}_i^{\text{Pos}}$	$\bar{E}_i^{\text{Neg}}$		$\bar{E}_i^{\text{Pos}}$	$\bar{E}_i^{\text{Neg}}$		$\bar{E}_i^{\text{Pos}}$	$\bar{E}_i^{\text{Neg}}$
A_1	0.0108	-0.0197	A_4	0.0253	-0.0451	A_7	0.0100	-0.0291
A_2	0.0173	-0.0105	A_5	0.0334	-0.0217	A_8	0.0342	-0.0216
A_3	0.0263	-0.0157	A_6	0.0576	-0.0213	A_9	0.0159	-0.0397

Table 3. The performance of each attribute.

Attribute	Performance	Attribute	Performance	Attribute	Performance
A_1	0.1973	A_4	0.2518	A_7	0.1834
A_2	0.3989	A_5	0.2291	A_8	0.1347
A_3	0.3322	A_6	0.2517	A_9	0.1198

Table 4. The values of Range_i and AI_i related to the nine attributes.

	Range_i	AI_i		Range_i	AI_i		Range_i	AI_i
A_1	0.0305	-0.3771	A_4	0.0704	-0.2833	A_7	0.0391	-0.6475
A_2	0.0278	0.3676	A_5	0.0551	0.2131	A_8	0.0558	0.2603
A_3	0.0420	0.2573	A_6	0.0789	0.4526	A_9	0.0556	-0.4146

obtained AI_i and Per_i, the AIPA plot can be drawn with AI_i on the vertical axis and Per_i on the horizontal axis, $i = 1, 2, ..., 9$, as shown in Figure 4. With the help of the AIPA plot, the influence of attributes concerning asymmetric effect and performance on CS could be interpreted [2].

It can be seen from Figure 4 that the attributes A_1, A_7 and A_9 are basic attributes whose requirements have not been fulfilled. The three attributes have great potential to lead to customer dissatisfaction, thus managers should give priority to these three attributes. The attributes A_5 and A_8 are excitement attributes with low performance. If performances of the two attributes are improved, then CS will increase significantly. A_4 is a basic attribute with high performance, which needs to be well-maintained. The attributes A_2, A_3 and A_6 are excitement attributes which have above average performance. The three attributes have a high potential in leading to CS. However, the three attributes do not give rise to dissatisfaction when their performances decrease. Thus, in allocating the limited resources of the hotel, the priority of the three attributes is relatively lower than other attributes.

CONCLUSION

In this paper, a method for conducting AIPA based on online reviews is proposed. Although online reviews have been used as the data source for several kinds of decision analysis, such as tourist destination analysis, product ranking, product recommendation, products improvement, brand analysis, consumer preference analysis, as far as we know, it is the first attempt to conduct AIPA based on online reviews. Comparing with the existing methods on conducting AIPA through surveys, the proposed method can obtain effective AIPA results of product and service attributes with lower cost and shorter time because products and services online reviews can be easily collected. The proposed method not only lays a good foundation for further conducting CS analysis and formulating products and services improvement strategies through online reviews, but also provides managers one more choice for conducting AIPA with lower cost and shorter time.

In terms of further research, since fake reviews may affect the accuracy of the proposed method, the fake reviews should be identified and removed beforehand to obtain more accuracy

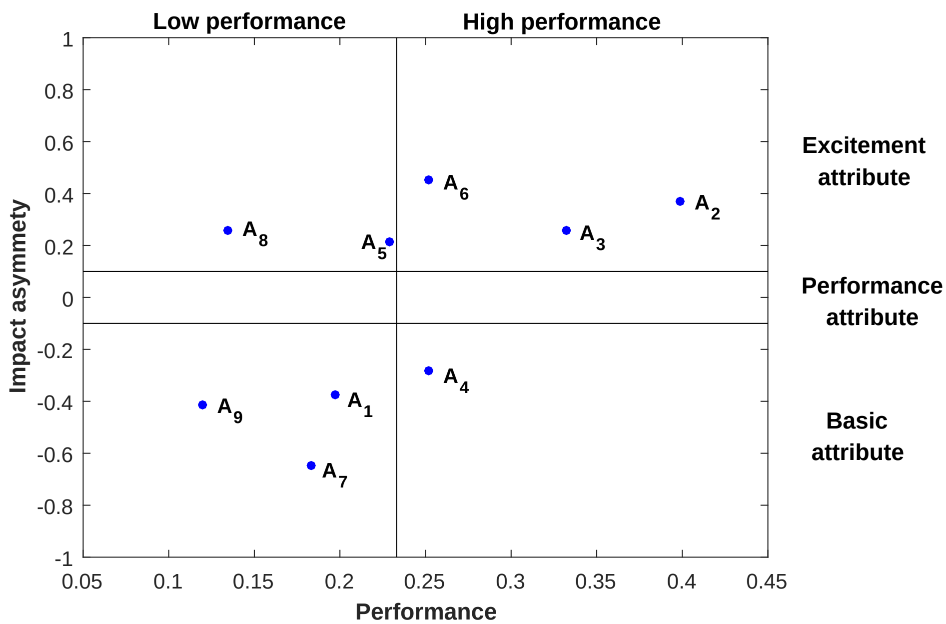

Figure 4. The AIPA plot of the five-star hotel related to the extracted nine attributes.

analysis results.

REFERENCES

1. J. A. Martilla and J. C. James, "Importance-performance analysis," *Journal of marketing*, vol. 41, no. 1, 1977, pp. 77–79.
2. M. Caber, T. Albayrak, and E. T. Loiacono, "The classification of extranet attributes in terms of their asymmetric influences on overall user satisfaction: An introduction to asymmetric impact-performance analysis," *Journal of Travel Research*, vol. 52, no. 1, 2013, pp. 106–116.
3. V. Mittal, W. T. Ross Jr, and P. M. Baldasare, "The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions," *Journal of Marketing*, vol. 62, no. 1, 1998, pp. 33–47.
4. J.-W. Bi et al., "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model," *International Journal of Production Research*, vol. 57, no. 22, 2019, pp. 7068–7088.
5. R. M. Groves, "Nonresponse rates and nonresponse bias in household surveys," *Public Opinion Quarterly*, vol. 70, no. 5, 2006, pp. 646–675.
6. F. Z. Xing, F. Pallucchini, and E. Cambria, "Cognitive-inspired domain adaptation of sentiment lexicons," *Information Processing & Management*, vol. 56, no. 3, 2019, pp. 554–564.
7. Q. Yang et al., "Segment-level joint topic-sentiment model for online review analysis," *IEEE Intelligent Systems*, vol. 34, no. 1, 2019, pp. 43–50.
8. J.-W. Bi et al., "Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews," *Tourism Management*, vol. 70, 2019, pp. 460–478.
9. E. Cambria et al., "SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings," *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018, pp. 1795–1802.
10. E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, 2016, pp. 102–107.
11. E. Cambria et al., "Sentiment analysis is a big suitcase," *IEEE Intelligent Systems*, vol. 32, no. 6, 2017, pp. 74–80.

ACKNOWLEDGEMENTS

This work was supported in part by the China Postdoctoral Science Foundation (Project No. 2019M661000), in part by the National Science Foundation of China (Project Nos. 71771043, 71871049 and 71971124), in part by Liaoning BaiQianWan Talents Program (Project No. 2016921027), in part by the Fundamental Research Funds for the Central Universities, China (Project No. N170605001) and the 111 Project

(B16009).

Jian-Wu Bi is currently a Research Associate (Postdoctoral) with College of Tourism and Service Management, Nankai University, Tianjin, China. His research interests include natural language processing, machine learning and their applications in management decision analysis. He received the Ph.D. degree in management science and engineering from Northeastern University, Shenyang, China, in 2019. Contact him at jwbi@nankai.edu.cn.

Yang Liu is currently a Professor with School of Business Administration, Northeastern University, Shenyang, China. His research interests lie in decision analysis and operation research. He received the Ph.D. degree in management science and engineering from Northeastern University in 2010. He is the corresponding author of this article. Contact him at liuy@mail.neu.edu.cn.

Zhi-Ping Fan is currently a Professor with School of Business Administration, Northeastern University, Shenyang, China. His research interests are decision analysis, operation research and knowledge management. He received the Ph.D. degree in control theory and applications from Northeastern University in 1996. Contact him at zpfan@mail.neu.edu.cn.