Understanding Customers' Insights Using Attribution Theory

Raksmey Sann, Pei-Chun Lai, Shu-Yi Liaw and Chi-Ting Chen

April 6, 2023

Understanding Customers' Insights Using Attribution Theory

ABSTRACT

This study attempts to detect associations between complaint attributions and specific consequences by guests of different star-rated hotels. A multifaceted approach is applied. First, a content analysis is conducted to transform textual complaints into categorically structured data. Then, an association rule technique is applied to discover potential relationships amongst complaint antecedents and consequences. Utilizing an Apriori rule-based machine learning algorithm, optimal priority rules for this study were determined for the respective complaining attributions for both the *antecedents* and *consequences*. Based on attribution theory, this study found that *Customer Service*, *Room Space and Miscellaneous Issues* received more attention from guests staying at *higher star-rated* hotels. Conversely, *Cleanliness* was a consideration more prevalent amongst guests staying at *lower star-rated* hotels. Practical implications are also discussed.

Keywords: Customer complaint, online review, attribution theory, Apriori algorithm

INTRODUCTION

To gain insights of online booking reviews to extant research on hotel guest profiles in the relationships of guest preferences, experiences and expectations, the big data and textual analysis methods are applied to analyse differences in models. In the hotel industry, for instance, by utilizing big data analysis generated from guest ratings, Liu, Teichert, Rossi, Li, and Hu (2017) identified the specific factors of hotel satisfaction and found that some lower star-rated hotels offer better perceived service quality than high star-rated hotels. Some studies revealed guest dissatisfaction were associated with cleanliness and facilities problem at lower star-rated accommodations. However, in higher star-rated hotels, more guest complaints were caused by service quality problems and issues related to high pricing (Fernandes & Fernandes, 2017, Hu et al., 2019). It highlights the possible link and the research gap between guest preferences towards individual hotel attributes and the hotel category for the enhancement of guest satisfaction and hotel competitive strengths. Therefore, this study investigates hotel guest preferences, experience and expectations via online guest review data. Incorporated with the application of attribution theory, the study further explores the relationships between complaint attributions and the consequences of guest behaviour from different hotel categories.

LITERATURE REVIEW

In machine learning, a commonly applied technique which is used to discover interesting relations in the data is known as ***Association Rules (ARs)*** (Martín et al., 2018). ARs is also referred to as **Shopping Cart** (or **Market Basket**) **Analytics**, and can be used to detect relationships or associations that exist between the specific values of categorical variables in large data sets (Hung, 2018a). ARs have been used to uncover hidden patterns. For instance: customers who order product A often also order product B and/or C; or employees who said positive things about initiative X also frequently complain about issue Y, but are happy with issue Z. For retailers, ARs can be used to predict which products consumers will potentially purchase or make suggestions for future purchases. For instance, Amazon.com and other retailers use ARs to make recommendations for similar products or items which are frequently purchased together (Kalgotra & Sharda, 2018). In this respect, the new knowledge which is produced via mining-ARs is considered as informative as the knowledge derived from classical ones (Jabbour et al., 2018). In this study, the ARs techniques is applied to explore the associated antecedents and consequences of online guest complaints from different star-rated hotels to provide the predication of specific subjects. For instance, a customer who complains about attribute A also complains about attribute B will happen in the hotel category X.

Discussing the differences between star-rated hotel guest preferences, in the study of Liu et al. (2017) shows that hotels of lower star designation are sometimes able to outperform hotels of higher designation in terms of the ratings that the guests assign. In terms of customer complaints; Ekiz, et al.(2012) pointed out "Rooms" received the highest number of complaints from luxury hotel guests, followed by Service Failures caused by inexperienced, unprofessional, or misbehaving staff. Fernandes and Fernandes (2017) analyzed guest reviews for hotels listed on TripAdvisor. Their findings revealed the categories of Cleanliness, Rooms, Sleep Quality, Bathrooms, Breakfast, and Facilities received the most negative statements for 2- to 3-star hotels. Comparatively, 4- to 5-star hotels received more negative statements directed at Customer Care, Location, and Value. These findings may indicate that core activities are central to service providers at budget hotels, while peripheral activities may enhance the value of the luxury hotels to meet the expectations of guests.

METHODOLOGY

Preparation and processing data

Step 1: collection and samples

This study collected the data from TripAdvisor site. Online reviews were collected with an overall rating of *one*- or *two*-star, as they considered as “complaint” (Sann et al., 2020). In total, we retrieved more than 2,000 samples from over than 300 hotels in United Kingdom.

Step 2: Coding procedures

First, every review was retried manually from the hotel webpages. Then content analysis was conducted in order to transform unstructured review data into structured data by following technique of Liu (2019). Finally, we identified 10 complaint attributes.

Step 3: Coding reliability

Percentage agreement was performed in order to guarantee the inter-coder reliability. This study followed the Sann et al. (2023) and Cenni and Goethals (2017). The inter-code reliability test was performed after coding of 5% and 10% of the full sets of total data. Both coding grids were higher than 90% which was in a good condition.

Apriori algorithm

An Apriori algorithm was applied in this study in order to identify the best rules for the model. A total of ten online complaining attributes were set to *input*, meaning they supplied the *antecedent* side, while the *Hotel Star-Ratings* were set to *target*, meaning they provide support for the *consequent*. A simple example of the complaining pattern sequence can be represented in the form of A → B, where A (called the *antecedent* or first attribute) and B (called the *consequent* or second attribute) are two disjoined item sets. This can be stated in the following form:

if *antecedent* (s) **then** *consequent* (s)

The minimum rule confidence threshold for those taken into consideration should be at least 70%, with a minimum antecedent support of 10% and a maximum number of four antecedents.

EMPIRICAL FINDINGS

The generated rules are presented based on a minimum confidence of 70% and maximum of four antecedents. Based on the analysis, the top seven priority rules for the different *star-rated* of hotels were compiled from the transactional data for the online complaining attributes (*Table 1*). The first rule listed in *Table 1* posits that if a guest makes an online complaint about *Miscellaneous Issues* and *Customer Service* but doesn't complain about *Safety* and *Location Accessibility*, then the guest stayed at a *higher star-rated* hotel. This rule has a confidence of 71.09%. This means that guests making complaints about *Miscellaneous Issues* and *Customer Service* stayed at *higher star-rated hotels* 71.09% of the time. Also, the support of the rule is 10.45% meaning that in the entire transactional database, if a guest made online complaints about *Miscellaneous Issues* and *Customer Service*, it is associated with *higher star-rated* hotels 10.45% of the time. In another example according to the 6th rule, if *Cleanliness = Yes* and *Customer Service = No* and *Room Space = No* and *Hotel Facility = No* then the confidence is 70.59% that the complaint was made by a customer of a *Lower Star-Rated* hotel. Therefore, if a customer makes an online complaint about *Cleanliness*, but not *Customer Service, Room Space, and Hotel Facility*, then in 70.59% of the time, the complaint applies to a *lower star-rated* hotel. Also, the support for this rule is 10.94%, meaning that in the entire transactional database, if a guest makes an online complaint about *Cleanliness*, but not about *Customer Service, Room Space, and Hotel Facility*, it will be associated with a *lower star-rated* hotel 10.94% of the time. *Table 1* demonstrates association rules for online complaining behavior.

Table 1: Association rules for online complaining behavior

Rule Ranking	**Consequent**	**Antecedent**	**Rule ID**	**Instances**	**Support %**	**Confidence %**	**Rule Support %**	**Lift**	**Deployability**
1	Hotel Star-Rating = Higher Star-Rated	Miscellaneous Issue = Yes and Customer Service = Yes and Safety = No and Location Accessibility = No	15	211	10.446	71.090	7.426	1.251	3.020
2	Hotel Star-Rating = Higher Star-Rated	Miscellaneous Issue = Yes and Customer Service = Yes and Safety = No	6	215	10.644	70.698	7.525	1.244	3.119
3	Hotel Star-Rating = Higher Star-Rated	Room Space = Yes and Cleanliness = No	1	320	15.842	70.625	11.188	1.243	4.653
4	Hotel Star-Rating = Higher Star-Rated	Customer Service = Yes and Cleanliness = No and Location Accessibility = No	10	752	37.228	70.612	26.287	1.242	10.941
5	Hotel Star-Rating = Higher Star-Rated	Customer Service = Yes and Miscellaneous Issue = No and Cleanliness = No and Location Accessibility = No	16	585	28.960	70.598	20.446	1.242	8.515
6	Hotel Star-Rating = Lower Star-Rated	Cleanliness = Yes and Customer Service = No and Room Space = No and Hotel Facility = No	13	221	10.941	70.588	7.723	1.635	3.218
7	Hotel Star-Rating = Lower Star-Rated	Cleanliness = Yes and Customer Service = No and Value for Money = No and Hotel Facility = No	14	231	11.436	70.563	8.069	1.635	3.366

CONCLUSIONS

The main purpose of this study is to detect associations between complaint attributions and specific consequences by guests of different star-rated hotels. From analysis, this study confirm that online complaint behavior differs among various star-rated hotels. Based on attribution theory, this study found that *Customer Service*, *Room Space and Miscellaneous Issues* received more attention from guests staying at *higher star-rated* hotels. Conversely, *Cleanliness* was a consideration more prevalent amongst guests staying at *lower star-rated* hotels.

Rule mining approach offers researchers and hotel managers to have a broad view of relationship that might occur between different complaint patterns. In addition, this technique offers a clear process and visualization that are easier to understand. Thus, hoteliers are recommending adopting association rule to assist identify issues.

References

Cenni, I., & Goethals, P. (2017). Negative hotel reviews on TripAdvisor: A cross-linguistic analysis. *Discourse Context & Media*, *16*, 22-30. https://doi.org/10.1016/j.dcm.2017.01.004

Fernandes, T. & Fernandes, F. (2017) Sharing dissatisfaction online: Analyzing the nature and predictors of hotel guests negative reviews. *Journal of Hospitality Marketing Management*, *27*, 127–150.

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism management*, *72*, 417-426. https://doi.org/10.1016/j.tourman.2019.01.002

Hung, C. (2018a). Association Rules: Sequence, Association and Link Analysis. In (pp. DM0018-DM0019). The Data-Shack Limited.

Jabbour, S., Mazouri, F. E. E., & Sais, L. (2018). Mining Negatives Association Rules Using Constraints. The First International Conference On Intelligent Computing in Data Sciences,

Kalgotra, P., & Sharda, R. (2018). BIARAM: A process for analyzing correlated brain regions using association rule mining. *Comput Methods Programs Biomed*, *162*, 99-108. https://doi.org/10.1016/j.cmpb.2018.05.001

Liu, J. W. (2019). Using big data database to construct new GFuzzy text mining and decision algorithm for targeting and classifying customers. *Computers & Industrial Engineering*, *128*, 1088-1095. https://doi.org/10.1016/j.cie.2018.04.003

Liu, Y., Teichert, T., Rossi, M., Li, H. X., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism management*, *59*, 554-563. https://doi.org/10.1016/j.tourman.2016.08.012

Martín, D., Martínez-Ballesteros, M., García-Gil, D., Alcalá-Fdez, J., Herrera, F., & Riquelme-Santos, J. C. (2018). MRQAR: A generic MapReduce framework to discover quantitative association rules in big data problems. *Knowledge-Based Systems*, *153*, 176-192. https://doi.org/10.1016/j.knosys.2018.04.037

Sann, R., Lai, P.-C., & Liaw, S.-Y. (2020). Online complaining behavior: Does cultural background and hotel class matter? *Journal of Hospitality and Tourism Management*, *43*, 80-90. https://doi.org/10.1016/j.jhtm.2020.02.004

Sann, R., Lai, P.-C., & Liaw, S.-Y. (2023). Understanding Customers' Insights Using Attribution Theory: A Text Mining and Rule-Based Machine Learning Two-Step Multifaceted Method. *Applied Sciences*, *13*(5), 3073. https://www.mdpi.com/2076-3417/13/5/3073