

University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Travel and Tourism Research Association: Advancing Tourism Research Globally

Exploring Destination Loyalty: Application of Social Media Analytics in a Nature-based Tourism Setting

Farshid Mirzaalian
University of Alberta

Elizabeth Halpenny
University of Alberta Faculty of Kinesiology, Sport, and Recreation

Follow this and additional works at: <https://scholarworks.umass.edu/ttra>

Mirzaalian, Farshid and Halpenny, Elizabeth, "Exploring Destination Loyalty: Application of Social Media Analytics in a Nature-based Tourism Setting" (2020). *Travel and Tourism Research Association: Advancing Tourism Research Globally*. 15.
https://scholarworks.umass.edu/ttra/2020/research_papers/15

This Event is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Travel and Tourism Research Association: Advancing Tourism Research Globally by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Exploring Destination Loyalty: Application of Social Media Analytics in a Nature-based Tourism Setting

Introduction

Loyalty is an important topic of study in hospitality and tourism research. Tourism providers understand the importance of loyal visitors, as their competitors offer similar attractions, services and experiences. Appreciation of how visitors form their destination loyalty and factors that influence their loyalty is important for the success of tourism destinations. User-generated content and electronic word of mouth posted by travelers online, provides a rich source of self-reported, publicly accessible, unconstrained data, enabling researchers to enter the minds of tourists without any set parameters and explore their thoughts on loyalty (Berezan, Raab, Tanford, & Kim, 2015). Obtaining market research data and understanding social interaction from online communities is an efficient and naturalistic method of collecting data. It also often outperforms traditional data collection methods such as focus groups and interviews, as people are more open and honest in expressing their perceptions online than in real life (Kozinets, 2002; Reid, 1996).

Analyzing sentiments and exploring themes of online reviews can help destinations understand the value of their brand in the minds of visitors, and whether the destination has been able to deliver its brand promise. Sentiment analysis of negative reviews, for instance, highlights where a destination has failed to deliver services that were claimed in its mission. On the other hand, analyzing the most enthusiastic reviews from loyal visitors can give destination marketing organizations an idea of why visitors were inspired to provide positive word of mouth and revisit intentions. The present study, which is exploratory in nature, presents a novel approach that applies different analytical techniques such as sentiment analysis, topic modeling, and text clustering to extract sentiments and topics of interest from tourists' conversational data on TripAdvisor from 2002 to 2019. It also identifies destination loyalty declarations using a keyword clustering approach. Previous destination loyalty literature was used to develop a keyword list to search for expressions of loyalty in online reviews. The robustness of loyalty clusters and optimal number of clusters were also assessed prior to final analysis. Four leading loyalty-focused categories of destination offerings were observed: glaciers, waterfalls, lakes and islands, and hiking and trails. Prioritization of visitor experience enhancements relating to these loyalty expressions inducing destination components are discussed.

Literature Review

To transform casual visitors to loyal patrons, destinations first need to know what visitors' expectations are, so that they can meet and potentially exceed those expectations by providing appealing services before, during and after their visit. Understanding how visitors form their destination loyalty and what factors influence their loyalty formation is important for the success of tourism destinations. There are three main approaches for defining and measuring tourist loyalty: measuring attitude, measuring behavior, or measuring a combination of two. The behavioral perspective focuses on a tourist's actual consumption behavior such as repeat visit duration, frequency and intensity (Oppermann, 2000). In contrast to the behavioral approach that produces only a static outcome of a dynamic process, the attitudinal perspective goes beyond and considers loyalty in terms of tourists' strength of affection toward a destination or attraction (Pritchard & Howard, 1997). Finally, a composite conceptualization of loyalty integrates both behavioral and attitudinal dimensions, by not only looking at the tourist's consumption behavior such as repeated

visits, but by considering future actions such as willingness to recommend to third parties (Oppermann, 2000), the strength of preference (Lee, Yoon, & Lee, 2007), and the feeling of attachment towards the place (Yuksel, Yuksel, & Bilim, 2010). Chen and Gursoy (2001) argued that a composite measure of loyalty (combination of both attitudinal and behavioral measures) provides the most accurate representation of destination loyalty. Identifying determinants of loyalty has been an important research topic among tourism researchers. While some loyalty-related researchers have focused on factors such as activity (Backman & Crompton, 1991), service quality (Baker & Crompton, 2000), and tourism providers (Morais, Dorsch, & Backman, 2004), other researchers have pointed out the importance of commitment to a specific place, or what is referred to as destination loyalty (Kyle, Graefe, Manning, & Bacon, 2004; Oppermann, 2000). The related destination loyalty literature background was applied to develop a keyword list that was used to detect destination loyalty expressions in online reviews. A corpus or collection of words that make up statements made by visitors posting on TripAdvisor about Jasper National Park was collated. These keywords included but not limited to: “revisit,” “visit again,” “come back,” “recommend,” and “worth.”

Methodology

All English travelers’ reviews about top natural attractions and park areas in Jasper National Park (JNP) were extracted from the third-party review website TripAdvisor, ranging from as early as December 2002 to October 2019. The reviews were collected in October 2019 (a total of 17224 English reviews). JNP is the largest national park in the Canadian Rockies and part of UNESCO’s Canadian Rocky Mountain Parks World Heritage Site (Parks Canada, 2019). Top natural attractions and park areas listed by TripAdvisor were as follows: Annette Lake, Athabasca Falls, Athabasca Glaciers, Columbia Icefield, Maligne Canyon, Maligne Lake, Mt. Edith Cavell, Mt. Edith Cavell Trail, Pyramid and Patricia Lakes, Spirit Island, Sulphur Skyline Trail, and Sunwapta Falls and Canyon.

The sentiment score was constructed by scoring the online reviews for positive and negative terms using Hu & Liu’s (2004) sentiment lexicon. Khoo & Johnkhan (2018) compared the effectiveness of different sentiment lexicons for sentiment categorization at the document level and sentence level. They reported that Hu & Liu’s (2004) lexicons outperform other lexicons for product review sentiment categorization, and obtain the best accuracy especially when a training corpus is not available (Khoo & Johnkhan, 2018). All characters were converted to lowercase in both training and test datasets. The sentiment score was calculated by adding a point for each positive word to the total score and deducting a point for each negative word (no points were given for neutral words) (Miner et al., 2012; Philander & Zhong, 2016). The Pos/Neg ratio score is computed as the ratio of overall positive lexicons in each location to overall negative lexicons, with any neutral lexicon discarded. The average number of words in any of 12 attractions were also reported, showing that despite significant differences between some attractions in terms of overall volume of reviews, all 12 locations are to some extent consistent in terms of average number of words used in online reviews.

The second approach, LDA topic modeling, was employed to effectively extract dimensions of the visitor experience from a large corpus of text data (Blei, Ng, & Jordan, 2003). LDA is the most common method for topic modelling and is a generalization of probabilistic latent semantic indexing (PLSI) (Hofmann, 1999). The LDA model was adopted instead of other text classification methods mainly because the LDA model not only surpasses other methods in

efficiently analyzing large-scale data at a highly granular level, but because it also helps to clarify the practical frequency of occurrence of each extracted dimension based on its intensity in the corpus (Guo, Barnes, & Jia, 2017). Revealed topics represent the important aspects related to tourists' experience and have a distribution across the online reviews depending on their frequency of occurrence.

Finally, a keyword clustering approach was implemented on the collection of online reviews that contained loyalty-expressed reviews. A term-document-matrix with the TF-IDF weighted review words was used for this clustering approach, where the matrix characterizes the basis of the K-means clustering algorithm with the cosine similarity as distance measure as recommended by the text-document-clustering literature (Huang, 2008; Menner, Höpken, Fuchs, & Lexhagen, 2016). Words with high TF-IDF values within a cluster then represent words often co-occurring in loyalty-expressed reviews and, thus, represent topics. The Elbow Method was used to examine the perplexity value (i.e., a measurement of how well a probability model predicts a sample), and to determine the appropriate number of topics for the LDA model (Ketchen & Shook, 1996).

Results

On average, there were 1435 online reviews per location, with Athabasca Falls having the most reviews (4319), and Annette Lake the fewest (87). Some locations with lower review volumes, such as Sulphur Skyline Trail and Spirit Island, appeared in the top five average sentiment score rank, as along with other well-known locations with higher review volumes such as Maligne Lake. Also, although the average sentiment score and the Pos/Neg ratio score ranks were aligned with one another for most of the locations, some attractions had meaningfully different ranks such as Sulphur Skyline and Mt. Edith Cavell Trails, Spirit Island, and Athabasca Falls. Another remarkable finding upon comparing TripAdvisor relative rank with sentiment and ratio scores is that lakes and islands are relatively ranked lower on TripAdvisor in contrast to higher sentiment and ratio ranks uncovered in our results. These attractions are Annette Lake, Pyramid and Patricia Lakes, Maligne Lake, and Spirit Island.

The LDA was applied to extract and label the dimensions of tourist experience across all collected online reviews from top touristic locations. The LDA identified 14 topics and within each topic showed the top-20 words and their relative weight. The labeling of dimensions was first conducted by one researcher and then confirmed by a second researcher. Labeling was based on the identification of a logical connection between the most frequent words for a topic. First five dominant topics were "Ice Walking," "Glacier Exploring Tour," "Scenic Waterfalls," "Water-based Activities," and "Waterfalls Visit Experience," respectively. Three of the dimensions represent tourists' perceptions of glaciers: "Ice Walking," "Glacier Exploring Tour," and "Glacier Visit Experience," while three dimensions correspond to tourists' hiking activities: "Trails and Pathways," "Hiking Activities," and "Forest Challenge" (see table 3). Other groups of dimensions represent lakes and islands (e.g., "Scenic Lakes and Islands," "Water/Cruise Tours," and "Water-based Activities") and waterfalls (e.g., "Scenic Waterfalls" and "Waterfalls Visit Experience"). The remaining dimensions show four distinct aspects of tourists' general experience (e.g., "Suggestions," "Weather," "General Experience").

After a detailed review of the hospitality and tourism loyalty literature by the authors, a vocabulary of destination loyalty keywords was developed. This keyword vocabulary was subsequently used to identify and separate loyalty-expressed reviews from the rest of the corpus.

A close reading check was applied to ensure the relatedness of the extracted reviews to destination loyalty conversations. Keyword clustering approaches including TF-IDF term document matrix and K-means clustering algorithm have been applied on the collection of destination loyalty online reviews. Results of the K-means clustering method along with the 4 themes of detected topics of LDA model suggest that destination loyalty expressions can be categorized into 4 main subjects, namely “waterfalls,” “glaciers,” “lakes and Islands,” and “hiking and trails.” Thus, all of the 12 touristic locations were further categorized and labeled into these 4 clusters based on the nature of the place and types of activities that take place in each (e.g., Athabasca Falls into waterfalls, Athabasca Glaciers into glaciers, Annette Lake into lakes and Islands, and Sulphur Skyline Trails into hiking and trails) (table 1). Top unigrams and bigrams of each category were identified. Bigrams are expected to improve the model performance by taking into consideration words that tend to appear together in the reviews associated with the 12 different locations. A thorough investigation of unigrams and bigrams within each category revealed prevalent characteristics and factors important to tourists when expressing their loyalty on social media.

Table 1. Suggested destination loyalty dimensions within 4 clusters through cross N-gram comparison.

| Glaciers | Waterfalls | Lakes & Islands | Hiking and Trails |
|----------------------------|--------------------|----------------------------|--------------------------|
| -Icefield Skywalk | -Natural wonder | -Cruise tours | -Challenging trail |
| -Ice Explorer tour | -Beautiful scenery | -Water-based activity | -Sense of accomplishment |
| -Icefield Sightseeing tour | -Short/Easy hike | -Nature photography | -Beautiful skyline |
| -Icefields | -Winter walk | -Wildlife viewing | |
| -Parkway mountain drive | | | |

Conclusion and Discussion

In this study, the sentiment analysis revealed that some touristic locations in JNP are outperforming others in terms of sentiment and ratio scores on social media, despite the fact that tourists less frequently reflect on their experiences at these places on social media, resulting in lower volumes of reviews (e.g., Sulphur Skyline Trail, Mt. Edith Cavell Trail, and Annette Lake). The presence of these less considered locations placed higher in the ranking suggests that average sentiment score can be a more informative measure than simple TripAdvisor rankings. While the average sentiment score and the Pos/Neg ratio score ranks were aligned with one another for most of the locations, a more detailed review of the scores reveals that some attractions had meaningfully different ranks such as Sulphur Skyline and Mt. Edith Cavell Trails, Spirit Island, and Athabasca Falls. Part of this difference in ranking can be explained by the fact that a higher number of neutral reviews with sentiment scores of zero reduce a location’s average score but has no effect on the Pos/Neg ratio score. This suggests that locations with a considerably higher on-average scores compared with their ratio scores may have subgroups of visitors with extremely strong feelings toward these locations (e.g., Sulphur Skyline and Mt. Edith Cavell Trails).

This study also proposes a novel approach to extract latent dimensions of tourist experience derived at a nature-based tourism destination, retrieved from online reviews. The relative

significance of the obtained dimensions is identified based on the intensity of the conversations around each. “Ice Walking,” “Glacier Exploring Tour,” “Scenic Waterfalls,” and “Water-based Activities” are the most important dimensions in our analysis. This supports the findings of prior studies that have proposed natural environment, beauty of the scenery, and glacier tours as key factors influencing tourism experience and destination image (Beerli & Martin, 2004; Purdie, 2013). Results of our LDA model strongly suggest that JNP tourism providers leverage destination management dimensions such as glacier and cruise tourism experiences (controlled factors). The quality of the interpretation provided by tour operators and improvement in both content and delivery techniques are crucial factors to optimize tourists’ tour experiences. Well planned interpretation more likely results in satisfying visit experiences for tourists, which in turn leads to positive word-of-mouth, recommendations, and repeat visitation (Hwang, Lee, & Chen, 2005). DMOs and tour operators can play an important role in filling the knowledge gaps through trainings and workshops, mentoring and internships, as well as providing information materials directly to tourists pre, during and post visit. This goal cannot be achieved without a clear and effective communication and liaison channels between DMOs and tour operators.

Another controlled dimension for destination managers was trails and pathways. Considering the exceptionally higher sentiment scores of trails and hiking locations, both from our results and on TripAdvisor rankings, improving infrastructure and informative aspects of hiking trails and pathways through strategic and operational plans for trail development is something that tourism providers should invest further in. DMOs should also understand the needs and characteristics of potential hikers, identify diverse constraints that prevent their trail use, and recognize factors that inspire and facilitate their use. DMOs can also develop partnerships across different public and private sectors to promote specific trail activities, hiking experiences and packages for target groups, for example through showcasing unique cultural, natural, and historical features of the trail.

This study advances investigations of destination loyalty through cluster analysis of TripAdvisor online reviews. After categorizing loyalty-expressed reviews into 4 clusters of glaciers, waterfalls, lakes and islands, and hiking and trails, top features within each cluster were presented and analysed. Our results revealed that different types of tours play an important role in recommendations and revisit intentions of JNP tourists (e.g., Columbia Icefield and Sightseeing tours, Glacier Skywalk, Maligne Lake Cruise tour). Water-based recreational activities such as kayaking and canoeing, boating, paddle boarding, and fishing were amongst highly recommended activities when visiting lakes and islands. Nature photography and wildlife viewing were other inspiring factors for destination loyalty expression in reviews. Aligned with our findings from sentiment analysis and topic modeling, hiking activities and trail attractions were notable motivators for tourists’ loyalty expressions on social media. Our results show that sense of accomplishment upon finishing longer hikes and more challenging trails together with beautiful skyline and alpine view are amongst reasons for sharing loyalty toward JNP online.

This study has several managerial implications. Tourism providers can not only verify underlying aspects of tourist experience from user-generated data but can also portray a perceptual mapping of touristic locations within their destination through a comprehensive analysis of online reviews. Moreover, there is a lack of understanding about the factors influencing destination loyalty in nature-based setting. Thus, this study enables DMOs to specify destination’s salient characteristics that influence tourists’ recommendations and revisits intentions. Our online review

analysis of JNP visitors reveals key dimensions of destination loyalty toward JNP, including informative and recreational tours, water-based recreational activities, and challenging trails.

References

- Backman, S. J., & Crompton, J. L. (1991). Differentiating between high, spurious, latent, and low loyalty participants in two leisure activities. *Journal of park and recreation administration*, 9(2), 1-17.
- Baker, D. A., & Crompton, J. L. (2000). Quality, satisfaction and behavioral intentions. *Annals of tourism research*, 27(3), 785-804.
- Berli, A., & Martin, J. D. (2004). Factors influencing destination image. *Annals of tourism research*, 31(3), 657-681.
- Berezan, O., Raab, C., Tanford, S., & Kim, Y. S. (2015). Evaluating loyalty constructs among hotel reward program members using eWOM. *Journal of Hospitality & Tourism Research*, 39(2), 198-224.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Chen, J. S., & Gursoy, D. (2001). An investigation of tourists' destination loyalty and preferences. *International Journal of Contemporary Hospitality Management*, 13(2), 79-85.
- Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, 59, 467-483.
- Hofmann, T. (2017, August). Probabilistic latent semantic indexing. In *ACM SIGIR Forum* (Vol. 51, No. 2, pp. 211-218). ACM.

- Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177). ACM.
- Huang, A. (2008, April). Similarity measures for text document clustering. In *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand* (Vol. 4, pp. 9-56).
- Hwang, S. N., Lee, C., & Chen, H. J. (2005). The relationship among tourists' involvement, place attachment and interpretation satisfaction in Taiwan's national parks. *Tourism Management*, 26(2), 143-156.
- Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. *Strategic management journal*, 17(6), 441-458.
- Khoo, C. S., & Johnkhan, S. B. (2018). Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *Journal of Information Science*, 44(4), 491-511.
- Kozinets, R. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, 39, 61-72.
- Kyle, G., Graefe, A., Manning, R., & Bacon, J. (2004). Predictors of behavioral loyalty among hikers along the Appalachian Trail. *Leisure Sciences*, 26(1), 99-118.
- Lee, C. K., Yoon, Y. S., & Lee, S. K. (2007). Investigating the relationships among perceived value, satisfaction, and recommendations: The case of the Korean DMZ. *Tourism management*, 28(1), 204-214.

- Menner, T., Höpken, W., Fuchs, M., & Lexhagen, M. (2016). Topic detection: identifying relevant topics in tourism reviews. In *Information and Communication Technologies in Tourism 2016* (pp. 411-423). Springer, Cham.
- Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.
- Morais, D. B., Dorsch, M. J., & Backman, S. J. (2004). Can tourism providers buy their customers' loyalty? Examining the influence of customer-provider investments on loyalty. *Journal of Travel Research*, 42(3), 235-243.
- Oppermann, M. (2000). Tourism destination loyalty. *Journal of travel research*, 39(1), 78-84.
- Parks Canada (2019). About Jasper National Park. <https://www.pc.gc.ca/en/pn-np/ab/jasper/> Accessed 11 November 2019.
- Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management*, 55(2016), 16-24.
- Pritchard, M. P., & Howard, D. R. (1997). The loyal traveler: Examining a typology of service patronage. *Journal of travel research*, 35(4), 2-10.
- Purdie, H. (2013). Glacier retreat and tourism: Insights from New Zealand. *Mountain Research and Development*, 33(4), 463-473.
- Reid, E. (1996). Informed consent in the study of on-line communities: A reflection on the effects of computer-mediated social research. *Information Society*, 12, 119-127.

Yuksel, A., Yuksel, F., & Bilim, Y. (2010). Destination attachment: Effects on customer satisfaction and cognitive, affective and conative loyalty. *Tourism management*, 31(2), 274-284.