Noname manuscript No.
(will be inserted by the editor)

Tourism Management through the Big Data Paradigm

Sonia DJEBALI · Guillaume GUERARD

Abstract This paper highlights the two-ways relationship between data and entities producing or generating these data in the context of tourism management. Indeed, tourists leave various traces of their planning and trip on social media and networks. Moreover, studying those data to feed tourism marketing also affects the future data left by tourists. Understanding the intrinsic and extrinsic relationship between entities should significantly improve the knowledge in this research field.

Keywords Tourism Management · Complex System · Big Data · Social complexity

1 Introduction

In today's world, tourism is one of the most prominent areas for the economy. It is considered one of the widest and fastest growing industries [5].

Tourism is a displacement phenomenon that fully participates in the global traffic of peoples, concepts, objects, imaginaries, experiences, norms and indeed, tourism represents a major vector of globalization, mobility and traffic [3]. Thanks to the increasing number of digital tools such as search engines, social media, the Internet of Things and mobile technologies, users tend to share their opinions about the activities, sites they have visited, and their travel storytelling, which has the effect that the tourism experience is no longer personal but collective.

The tourism industry faces big data challenges, i.e. enormous volume and heterogeneous data. There has not existed a uniform definition of big data. A famous definition is the $3V$, which characterizes big data as Volume, Variety and Velocity [8]. This definition is extended to $11V$ [11].

Léonard de Vinci Pôle Universitaire, Research Center,
92 916 Paris La Défense, France
E-mail: FirstName.SecondName@devinci.fr

In this paper, we will focus on three V, Value, Variability, Veracity which allude to complex system characteristics as Emergence, Evolution and Environment. The Value is concerned mainly with extracting value and knowledge from data; the Variability is the nature of changing, shifting, mutation, and modification abruptly or with time and the Veracity is accuracy, truthfulness and meaningfulness of data.

Big data allowed a more accurate understanding of tourism demand, tourist behavior, tourist satisfaction and other tourism issues. Therefore, the complexity of interactions between digital tools and tourism management is mostly studied in one way without including the notion of feedbacks between those two systems.

2 Big Data and Tourism Research

In the literature, many research fields in tourism management exist as tourism recommendation, tourism marketing, tourism profiling, etc [9]. With the support of big data, the way to understand and analyze tourism management has evolved due to new sources of data. Those sources are grouped into three categories: *users* activities on social media and social networks, their *devices*, and *operations* or the digital traces left on the web. The following list refers to the sources and the linked research fields:

1. Users: they are people who use social media and social networks as Instagram, Flickr, TripAdvisor and left online textual or photo data [14]. Those data contain many metadata in the photo like GPS position, hashtag, temporal data. The drawback of those data is the reliability like fake reviews [12] or temporal gap between the visits and the reviews.
   - *Online textual data:* from textual data, scientists extract tourist sentiments and tourism recommendation.
   - *Online photo data:* those data are used to analyze tourist behaviors, tourism marketing and tourism recommendation.
2. Devices: they include smart phones and sensors. Smart phones provide information about the GPS localization, RFID transactions; sensors give details about meteorological data [13].
   - *Mobile GPS and RFID data:* those data are useful for tourists spatial-temporal behavior analysis, pattern extraction, tourism recommendation [2], tourism flow forecast [15]. However, the privacy concerns and the data processing may create a data graveyard due to incompatibility or unexploitable database.
   - *Meteorological data:* the weather factors have an impact on tourism destinations or attractions. The behavior analysis, tourism recommendation and tourism flow differ depending of the meteorological data [7].
3. Operations: they refer to data left by users when searching on the web, Web page, online booking, etc. [10]. They are analyzed for tourism demand prediction, tourist behavior analysis and tourism marketing. As devices

data, the question of privacy and estimation biases may produce a data graveyard.

All studies presented above are frequently conducted in one perspective, how to influence the tourism industry from the knowledge extracted from the big data. However, tourists are at the origin of data. They also affect the information retrieved. In this manner, there exists a two-ways interaction (or bidirectional interaction) between tourists and digital tools which have to be studied.

3 Tourism as a Complex System

New issues are emerging because the way of enjoying tourism has changed with the rise of social media and networks. Many hypotheses and problems born when the data are questioning about the social intrinsic impact i.e. interaction between entities. They can be characterized through a complex system model criterion as follows:

1. Feedbacks loop and environment: social media and networks influence the behaviors of tourists. But how tourists' behaviors impact the latter? How does it affect the flow of tourists, tourist spending, etc.? How to define and understand the impact of positive and negative feedback?
2. Emergence: social media Influencers or tourist operators are changing the demand [4]. How to define the targeted tourists, how the economy and tourist behaviors are impacted? Are there trends, seasonality, punctual events? Can we define in time and space the tourist impact of the emergence of new behaviors? How can we define the effects of crowds? Is it relevant and suitable to determine profiles of tourists in terms of destinations, attractions, spending, etc.? with what degree of data entanglement?
3. Rule system: tourists receive information from various media. Their decisions about where to go and what to visit depends internally on their personality, and externally on the influence of the media and their social network relationships. How to model these complex interactions between heterogeneous elements? How to extract from a model the relevant information about the studied interaction (e.g. define an indicator of tourist mobility in a region using TripAdvisor [6])?
4. Chaos and Evolution: is tourism stable in terms of destination and places to visit? Are there points of convergence like overtourism [1]? How does tourism evolve over time and space in a socio-economic context?

4 Conclusion

Through current tourism management studies, the information extracted from Big Data is done in a one-way, omitting the two-way impact between tourists and data. Investigating this impact through the prism of a complex system

model should considerably improve the understanding of this research field as suggested in this paper.

References

1. Alonso-Almeida, M.d.M., Borrajo-Millán, F., Yi, L., et al.: Are social media data pushing overtourism? the case of barcelona and chinese tourists. Sustainability **11**(12), 3356 (2019)
2. Baccar, L.B., Djebali, S., Guérard, G.: Tourist's tour prediction by sequential data mining approach. In: International Conference on Advanced Data Mining and Applications, pp. 681–695. Springer (2019)
3. Chapuis, A., Boukhris, L.: L'Espace Politique **28** (2016). URL https://hal-upec-upem.archives-ouvertes.fr/hal-01398998
4. Chatzigeorgiou, C.: Modelling the impact of social media influencers on behavioural intentions of millennials: The case of tourism in rural areas in greece. Journal of Tourism, Heritage & Services Marketing (JTHSM) **3**(2), 25–29 (2017)
5. Cooper, C., Hall, C.M.: Contemporary tourism. Routledge (2007)
6. Djebali, S., Loas, N., Travers, N.: Indicators for measuring tourist mobility. In: International Conference on Web Information Systems Engineering, pp. 398–413. Springer (2020)
7. Kim, S., Park, J.H., Lee, D.K., Son, Y.h., Yoon, H., Kim, S., Yun, H.J.: The impacts of weather on tourist satisfaction and revisit intention: a study of south korean domestic tourism. Asia Pacific Journal of Tourism Research **22**(9), 895–908 (2017)
8. Laney, D., et al.: 3d data management: Controlling data volume, velocity and variety. META group research note **6**(70), 1 (2001)
9. Li, J., Xu, L., Tang, L., Wang, S., Li, L.: Big data in tourism research: A literature review. Tourism Management **68**, 301–323 (2018)
10. Pan, B., Xiang, Z., Law, R., Fesenmaier, D.R.: The dynamics of search engine marketing for tourist destinations. Journal of Travel Research **50**(4), 365–377 (2011)
11. Patgiri, R., Ahmed, A.: Big data: The v's of the game changer paradigm. In: 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), pp. 17–24. IEEE (2016)
12. Schuckert, M., Liu, X., Law, R.: Hospitality and tourism online reviews: Recent trends and future directions. Journal of Travel & Tourism Marketing **32**(5), 608–621 (2015)
13. Shoval, N., Ahas, R.: The use of tracking technologies in tourism research: the first decade. Tourism Geographies **18**(5), 587–606 (2016)
14. Zeng, B., Gerritsen, R.: What do we know about social media in tourism? a review. Tourism management perspectives **10**, 27–36 (2014)
15. Zheng, W., Huang, X., Li, Y.: Understanding the tourist mobility using gps: Where is the next place? Tourism Management **59**, 267–280 (2017)