Recommendation of Tourism Resources Supported by Crowdsourcing

Fátima Leal[1,3], Benedita Malheiro[2,3] and Juan Carlos Burguillo[1]

[1]School of Telecommunications Engineering,University of Vigo, Spain
E-mail: fatimaleal2@gmail.com and J.C.Burguillo@uvigo.es

[2]School of Engineering, Polytechnic Institute of Porto, Portugal

[3]INESC TEC, Porto, Portugal
E-mail: mbm@isep.ipp.pt

Abstract

Context-aware recommendation of personalised tourism resources is possible because of personal mobile devices and powerful data filtering algorithms. The devices contribute with computing capabilities, on board sensors, ubiquitous Internet access and continuous user monitoring, whereas the filtering algorithms provide the ability to match the profile (interests and the context) of the tourist against a large knowledge bases of tourism resources. While, in terms of technology, personal mobile devices can gather user-related information, including the user context and access multiple data sources, the creation and maintenance of an updated knowledge base of tourism-related resources requires a collaborative approach due to the heterogeneity, volume and dynamic nature of the resources. The current PhD thesis aims to contribute to the solution of this problem by adopting a Crowdsourcing approach for the collaborative maintenance of the knowledge base of resources, Trust and Reputation for the validation of uploaded resources as well as publishers, Big Data for user profiling and context-aware filtering algorithms for the personalised recommendation of tourism resources.

Keywords: Crowdsourcing, Big Data, Trust and Reputation, Recommendation.

1 Problem Definition

Tourism is an enriching activity which promotes relaxation and well-being. When faced with the prospect of a travel, a tourist searches for a solution compatible with his tastes, interests and context, including time and money constraints. However, tourism resources are highly heterogeneous and are interdependent such as transportation, food, accommodation and attractions. This scenario comprises the modelling of users and resources - Profiling - and the matching of resources to users - Recommendation.

Travelling changed dramatically with mobile devices. These devices contribute with computing capabilities, on board sensors, ubiquitous Internet access, continuous user monitoring, collecting large volumes of tourist-related data (personal, context-aware and online-shared data). These features support the enrichment of the tourist experience during travelling and allow sharing relevant data in near real time. In this scenario, the Big Data concept emerges, encompassing both user modelling – the building of the user profiles from the gathered user-related information (personal, context-aware and online-shared data) – and tourism resources modelling (data regarding points of interests, restaurants, accommodation, *etc.*).

The research problem is focused on the heterogeneity of tourism resource, regarding its nature (different categories), lifespan (ephemeral or long lasting) and number. Ephemeral resources, *e.g.*, seasonal offers, are dynamic and new resources are constantly emerging. To allow the continuous building, updating and refinement of the tourism resource knowledge base, the system relies on stakeholder (end-users and businesses) inputs. This Crowdsourcing of tourism-related data, which takes advantage of the current data sharing trend on social networks, relies on Trust and Reputation models of the publishers for content validation. While, in terms of technology, personal mobile devices can gather user-related information, including the user context and access multiple data sources, the creation and maintenance of an updated knowledge base of tourism-related resources requires a collaborative approach due to the heterogeneity, volume and dynamic nature of the resources.

Finally, the matching of tourism resources according to the tourist profile is a complex challenge due to the size of the resource search space and of the tourist-related data.

We intend to address the problem by designing a Tourism Recommendation System supported by Crowdsourcing. The system includes a Profiler, a Crowdsourcer with Content Validation and a Recommender and adopts, in terms of implementation, cloud and agent-based computing paradigms. Thus, the current PhD thesis aims to contribute to the: (*i*) collaborative maintenance of the tourism resource knowledge base through Crowdsourcing; (*ii*) the validation of uploaded resources through the creation of Trust and Reputation models of publishers; (*iii*) the construction of user profiles supported by Big Data techniques; and (*iv*) Personalised Context-aware tourism resource Recommendation.

2   Literature Review

For countries, tourism is a source of revenues and contributes to the development of the local economy. Tourism, as a business, targets multiple segments of the public, offering specialised packages, and ranges an endless variety of resources and offers. Technology introduced a remarkable change in the tourism business model since, nowadays, people tend to rely more on Web resources than on standard travel agencies to obtain information, plan and book travels. However, the diversity, amount, dependency and volatility of tourism-related offers, the complexity of the task and the time required to find a solution, satisfying the context, interests and constraints of the potential tourist, makes this problem unmanageable for humans. In this scenario and taking into account the research problem of this PhD thesis, it is important to survey four different areas: Big Data, Crowdsourcing, Trust and Reputation and Context-aware Recommendation Systems.

2.1 Big Data

The concept of Big Data refers to the processing of large amounts of data stored disorderly over time. The volume of information nowadays available regarding tourism resources and tourist experiences, due to continuous data sharing, falls into the Big Data category. Akerkar, R. (2012) states that Big Data allows companies to

create better products and services by gathering information from numerous external sources, *e.g.*, travel agency portals, carriers or social networks.

Chareyron *et al.* (2014) and Akerkar (2012) identify Big Data as a new challenge for tourism. The World Travel & Tourism Council (WTTC) & Future Foundation (2014) discuss the effects of Big Data in travel industry, including the perils and advantages of using of Big Data. In terms of user profiling, we intend, like Pan, B., & Yang, Y. (2015), to combine multiple data sources through Big Data to improve the accuracy of recommendations. Other examples of tourism applications supported by Big Data include Chareyron *et al.* (2014) and Fuchs *et al.* (2014). While the first application analyses and detects typical and atypical tourist behaviour in heterogeneous social networks so that economic stakeholders can address the needs of tourists, the latter, which regards a Swedish touristic destination, performs Big Data analytics on a knowledge infrastructure framework gathered from different sources.

2.2 Crowdsourcing

The Crowdsourcing concept was introduced in 2006 by Jeff Howe (Howe, 2006). In general, Crowdsourcing is a process of getting work done by a crowd of people, *i.e.*, it corresponds to any collective and collaborative activity performed by a large number of volunteers via information and communication technologies. Few Crowdsourcing based approaches are found in the technology enhanced tourism domain literature. Within the present PhD thesis, Crowdsourcing is intended to perform the collaborative updating of the tourism resources knowledge base, using as workers the tourists and tourism businesses. In our perspective, Crowdsourcing can be explored using the main tourism stakeholders: businesses and tourists. According to Sigala (2015), tourism businesses regard the tourist crowd know-how as a valuable contribution for personalised marketing. In addition, Crowdsourcing, while a continuous source of tourist-generated, tourist-shared and tourist-updated data, promotes intangible tourism experiences (Sigala & Gretzel, 2014). The reunion of all tourism stakeholders in a single Crowdsourcing system promotes the gathering and updating of tourism knowledge. On the one hand, businesses will be able to define and recommend new packages based on the discovery of new tourism trends and resources and, on the other hand, the tourists will receive a wider and updated range of personalised relevant recommendations, contributing to improved travelling experiences

2.2.1 Crowdsourcing in Tourism Systems

Web-based tourism Crowdsourcing typically takes the form of collaborative maps, *e.g.* OpenStreetMap, wikis, *e.g.* Wikivoyage, or is incorporated in dedicated Web applications. OpenStreetMap (www.openstreetmap.org [Nov. 23, 2015]) is a map built by volunteers who collect and share map data such as streets, forests, rivers, lakes, *etc*. Wikivoyage (www.wikivoyage.org [Nov. 23, 2015]) is a collaborative wiki, similar to Wikipedia, dedicated solely to tourism. The Personalised Travel System based on Crowdsourcing by Zhuang *et al.* (2014) and TagTagCity (www.tagtagcity.com [Nov. 23, 2015]) are instances of tourism Web applications supported by Crowdsourcing. While in the first users search, recommend and update existing travel-related contents, TagTagCity, which is a Belgium Web-based system, using a map-based search engine to crowd-source data regarding a location.

Mobile tourism applications rely on Crowdsourcing to collect data for future and improved recommendations. Tiwari & Kaushik (2014, 2015) present a mobile location-aware tourism recommendation system enriched by Crowdsourcing. The crowd-sourced information is obtained through location-related questionnaires and is used to enrich the recommendations. Kamino (www.gokamino.com [Nov. 23, 2015]) is a location-based mobile application which offers crowd-sourced city guides, including attractions, restaurants, shops and other relevant complimentary information, while the user is walking around the city. Stereopublic (www.stereopublic.net [Nov. 23, 2015]) is a mobile Crowdsourcing application which gathers and shares data regarding quiet public spaces. The crowd adds and shares places on social networks together with audio clips and images. Identically, we intend to develop a mobile application to gather automatically context-aware data and crowd-source new resource data, enriching the system's knowledge base.

Crowd-sourced approaches have also been adopted by travel companies in order to find improved solution for individuals and groups of tourists. The HolidayCheck (www.holidaycheck.com [Nov. 23, 2015]) travel platform, which works with several well-known tour operators, relies on the crowd to rate hotels, share photos, videos and travel experiences. Rating, sharing and posting correspond to the type of user feedback and crowd-sourced data we are considering.

In terms of back-end systems, Vohnout *et al.* (2014) present SmartTouristData, a proposal for the integration of open linked tourism data sources. In this case, the Crowd-sourced data is provided by Linked Open Data (LOD) repositories such as OpenStreetMap, Open Weather Map or Wikitravel. Similarly, we intend to use Wikivoyage as a repository of crowd-sourced information. Finally, Bachrach *et al.* (2014) propose Crowdsourcing as a means to generate the initial data set for a personalised tourist attraction recommendation system. In the first stage, they generate a catalog of potential attractions in a target destination, where each entry is tagged with the attraction features, and, then, in the second stage, they use the Amazon Mechanical Turk (AMT) Crowdsourcing platform to ask users to validate and tag entries with the attraction features. If we do not find an appropriate data set for evaluation, we may adopt an identical approach to build our initial data set.

2.3 Trust and Reputation

Trust and Reputation systems represent a significant trend in decision support when choosing tourism resources. Trust and Reputation are distinct, but intrinsically linked concepts, *e.g.*, "*I trust that hotel because it has good reputation*". Jøsang *et al.* (2007) identify the different interpretations of Trust and Reputation, depending on the situation. In context of this PhD thesis, Trust defines the reliability of workers and resources based on direct experience, while Reputation is based on third party experiences, *e.g.*, the crowd. Regarding the building and application of Trust and Reputation models of the users to validate crowd-sourced tourism information, Marchiori *et al.* (2010) present a destination reputation model based on the RepTrak model. RepTrak is a corporate reputation model designed by the Reputation Institute (www.reputationinstitute.com [Nov. 23, 2015]) to measure and track general public and multi-stakeholder perceptions about companies.

2.4 Mobile Context-aware Recommendation Systems

Recommendation Systems are tools which produce personalised recommendations before a large variety of choices taking into account the current user behaviour. There are three generic recommendation methodologies: (*i*) Content-based; (*ii*) Collaborative; and (*iii*) Hybrid. In our case, we intent to develop a context-aware hybrid approach.

In the context-aware tourism mobile applications literature, Borràs *et al.* (2014) present a survey on tourist applications (mobile and Web-based) grouping the different technologies. Gavalas *et al.* (2014) detail the functions commonly offered by existing mobile tourism Recommender Systems (RS) prototypes. In turn, Luz *et al.* (2010) and Chen & Kotz (2000) focus on older applications.

3 Conceptual Development

The literature review on information and communication technologies (ICT) in tourism domain identifies important challenges to be explored in order to improve the tourist experience and the tourism business. In the context of this research, the users include tourism operators, such as travel agencies, hotels, restaurants, etc., and tourists. A user can play the role of a worker, when he publishes or classifies contents, and of a tourist, when they receive and explore recommendations. Workers can upload, classify and share content, building a shared and crowd-sourced repository of tourist resources. Figure 1 presents the conceptual framework illustrating the data flows between users and system components.
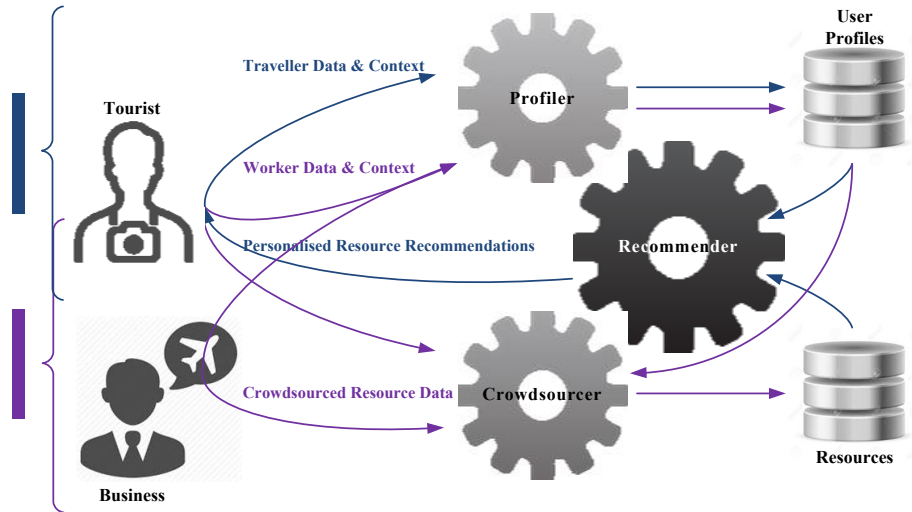
Figure 1. Conceptual Framework

In traveller mode, the mobile tourism application provides the tourist context and personal data, receives personalised resource recommendations and reports all feedback (clicks, rates and likes). In worker mode, the mobile tourism application facilitates the submission of the tourist and/or tourism business (context,

personal/institutional) data as well as of the crowd-sourced (posts, shares and comments) resource data. The Profiler builds user profiles based on the context, personal data and perceived Trust and Reputation. The Recommendation System matches traveller profiles with existing resources, taking into account the context, type, rating and credibility of the resources, *i.e.*, the trust and reputation of the publishers/workers.

4 Proposed Methodology

This PhD research is focused on Big Data, Recommendation and Crowdsourcing as well as on the exploration of Trust and Reputation mechanisms for content validation and publisher/worker modelling.

In the tourism domain it is fundamental to know the actual behaviour of tourists, including the information and communication technologies that they use before, during and after travelling. This study can be done through surveys which gather relevant information about tourist's behaviour and opinions. We created and distributed a questionnaire among tourism teachers, students and researchers, travel agencies and tourism enthusiasts (the general public). The results of this tourism survey will influence the selection of the information sources to use in the following PhD tasks, namely, in building and modelling of the user profile and the tourism resources.

The definition and building of the models of users (profiling) and the resources (recommendation) will be based on the crowd-sourced data as well as on the user data collected by mobile devices. This task involves Big Data and Recommendation techniques. In terms of tourism businesses, the Big Data approach will help to discover important travel patterns and trends, empowering them with the ability to enhance and improve the travel experiences of customers.

Similarly, the Crowdsourcing will support the maintenance of the system's knowledge base enriching both tourist and business activities. We plan to explore this topic with the adoption of WikiVoyage - the Wikipedia for tourism.

Trust and Reputation models of the workers will be built to validate the crowd-sourced content and, thus, improve recommendations. The trustworthiness and reputation of a worker depends on the quality of the volunteered contents and feedback. While trust is a direct relationship between users, reputation is based on the community perspective. A prolific worker with widely accepted and rated contributions will have a high reputation within the community. A user, which accepts and rates highly recommendations regarding resources uploaded by a given worker, establishes a strong trust relationship with the worker. This PhD thesis aims to design new or improved Trust and Reputation algorithms to explore these relationships in order to validate the crowd-sourced content.

The platform execution and the persistent data storage will rely on cloud computing resources. The resulting prototype system, which encompasses the different research topics, should be based on open source technologies.

5   Anticipated Results

The current PhD thesis aims to enrich the traveller experience by supporting the creation of a collaborative knowledge base of tourism resources and providing improved context-aware personalised suggestions. In terms of contributions, we intend to apply: (*i*) Trust and Reputation to validate crowd-sourced data regarding tourism resources as well as to model publishers; (*ii*) improve Big Data filtering algorithms for user (tourist and tourism business) profiling; and (*iii*) personalised context-aware resource Recommendation. As far as we know, this combination of information Crowdsourcing (validated by Trust and Reputation), Profiling (supported by Big Data algorithms) and Recommendation (assisted by Trust and Reputation) is a novel approach for mobile tourism applications.

We anticipate improvements on the personalised context-aware recommendations provided to tourists because: (*i*) user profiles will encompass both worker and tourist behaviours; (*ii*) the influence of irrelevant, incorrect or outdated resources as well as biased publishers will be reduced by the application of the Trust and Reputation models; and (*iii*) the knowledge base will be continuously updated.

In terms of evaluation, we intend to perform both offline and online assessment. For the offline tests, we are currently looking for an appropriate data set. Then, we intend to explore and use online Wikivoyage as Crowdsourcer together with the Trust and Reputation modelling. In the end, we plan to create and distribute a questionnaire among users (including the ones that participated in the initial questionnaire) and, thus, collect the users opinion.

References

Akerkar, R. (2012). Big Data & Tourism (12). Kolhapur, India: TMRF.

Bachrach, Y., Ceppi, S., Kash, I. A., Key, P., Radlinski, F., Porat, E., Armstrong, M., & Sharma, V. (2014). Building a personalized tourist attraction recommender system using crowdsourcing. In *Proceedings of the 2014 International Conference on Autonomous Cgents and Multi-Agent Systems* (pp. 1631-1632). International Foundation for Autonomous Agents and Multiagent Systems.

Borràs, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, *41*(16), 7370-7389.

Chareyron, G., Da-Rugna, J., & Raimbault, T. (2014). Big Data: a new challenge for tourism. In *Big Data (Big Data), 2014 IEEE International Conference on* (pp. 5-7). IEEE.

Fuchs, M., Höpken, W., & Lexhagen, M. (2014). Big data analytics for knowledge generation in tourism destinations–A case from Sweden. *Journal of Destination Marketing & Management*, *3*(4), 198-209.

Gavalas, D., Konstantopoulos, C., Mastakas, K., & Pantziou, G. (2014). Mobile recommender systems in tourism. *Journal of Network and Computer Applications*, *39*, 319-333.

Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine*, *14*(6), 1-4.

World Travel and Tourism Council & Future Foundation. (2014). Big Data: Insights for Travel & Tourism. London, UK: Future Foundation.

Jøsang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision support systems*, *43*(2), 618-644.

Marchiori, E., Inversini, A., Cantoni, L., & Dedekind, C. (2010). Towards a tourism destination reputation model. A first step. In *Proceedings of the 6th International Conference*

"Thought Leaders in Brand Management", (Lugano, Switzerland, 18-20 April 2010), CD-ROM (ISBN: 978-88-6101-006-2) (pp. 921-930).

Pan, B., & Yang, Y. (2015). Book Chapter for: Management Science in Hospitality and Tourism: Theory, Practice and Applications.

Sigala, M., Christou, E., & Gretzel, U. (Eds.). (2012). *Social media in travel, tourism and hospitality: Theory, practice and cases.* Ashgate Publishing, Ltd.

Sigala, M. (2015). Gamification for Crowdsourcing Marketing Practices: Applications and Benefits in Tourism. In *Advances in Crowdsourcing* (pp. 129-145). Springer International Publishing.

Tiwari, S., & Kaushik, S. (2014). Information Enrichment for Tourist Spot Recommender System Using Location Aware Crowdsourcing. In *Mobile Data Management* (MDM), *2014 IEEE 15th International Conference on* (Vol. 2, pp. 11-14). IEEE.

Tiwari, S., & Kaushik, S. (2015). Crowdsourcing Based Fuzzy Information Enrichment of Tourist Spot Recommender Systems. In *Computational Science and Its Applications –* ICCSA 2015 (pp. 559-574). Springer International Publishing.

Vohnout, P., Cerba, O., Kafka, S., Fryml, J., Krivanek, Z., & Holy, S. (2014). SmartTouristData approach for connecting local and global tourist information systems. In *IST-Africa Conference Proceedings*, 2014 (pp. 1-6). IEEE.

Zhuang, Y., Zhuge, F., Chiu, D. K., Ju, C., & Jiang, B. (2014). A Personalized Travel System Based on Crowdsourcing Model. In *Advanced Data Mining and Applications* (pp. 163-174). Springer International Publishing.

Acknowledgements