

Towards a social and context-aware mobile recommendation system for tourism

Ricardo Colomo-Palacios^{a,*}, Francisco José García-Peñalvo^b,
Vladimir Stantchev^c, Sanjay Misra^d

^a Faculty of Computer Sciences, Østfold University College, Postboks 700, 1757 Halden, Norway

^b Department of Computer Sciences, University of Salamanca, Plaza de los Caídos s/n, 37008 Salamanca, Spain

^c Institute of Information Systems, SRH University Berlin, Ernst-Reuter-Platz 10, 10587 Berlin, Germany

^d Department of Computer Engineering, Atilim University, Kizilcasar Mahallesi, Incek 06836, Ankara, Turkey

A B S T R A C T

Loyalty in tourism is one of the main concerns for tourist organizations and researchers alike. Recently, technology in general and CRM and social networks in particular have been identified as important enablers for loyalty in tourism. This paper presents POST-VIA 360, a platform devoted to support the whole life-cycle of tourism loyalty after the first visit. The system is designed to collect data from the initial visit by means of pervasive approaches. Once data is analysed, POST-VIA 360 produces accurate after visit data and, once returned, is able to offer relevant recommendations based on positioning and bio-inspired recommender systems. To validate the system, a case study comparing recommendations from the POST-VIA 360 and a group of experts was conducted. Results show that the accuracy of system's recommendations is remarkable compared to previous efforts in the field.

Keywords:

Tourism
Customer relationship management
Bio-inspired algorithms
Pervasive
Geographic information systems

1. Introduction

Tourism is one of the leading sectors in the world. According to the United Nations World Tourism Organization in its report *Tourism Highlights 2014 Edition*, tourism reaches 9% of the GDP (including direct, indirect and induced impact) and employs 1 out of every 11 persons worldwide [1]. Moreover, according to this report, over the past six decades, tourism has experienced continued expansion and diversification, becoming one of the largest and fastest-growing economic sectors in the world. Many new destinations have emerged in addition to the traditional favourites of Europe and North America.

Focusing on Europe, according to Tourism statistics from Eurostat [2], one in ten enterprises (2.2 million companies) in the European non-financial business economy belongs to the tourism industries. These organizations employ around 12 million persons. Tourism industries account for 3.6% of the turnover and 5.5% of the value added of the non-financial business economy. According to the Flash Eurobarometer 414, Preferences of Europeans towards tourism [3], published by the European Commission, Spain is again the most popular foreign destination in 2014, visited by 15% of respondents from other countries, followed by Italy (11%), France (10%) and Germany (8%). The next most popular destinations are Austria (6%), North Africa or the Middle East (6%), Greece (5%), and Asia or Oceania (5%). However, Spain, according to [2], is also very popular with non-EU residents. Spain was the most common tourism destination in the EU for non-residents, reaching

252 million nights spent in tourist accommodation establishments, or 21.2% of the EU-28 total. Across the EU, Spain is followed by Italy (185 million nights) and France (131 million nights), which together accounted for nearly half (47.7%) of the total nights spent by non-residents in the EU-28. More specifically, according to World Travel & Tourism Council [4], by 2013 the direct contribution of Travel & Tourism to Spanish GDP was 5.7% and the total contribution was 15.7%; these facts lead to a total contribution of Travel & Tourism to employment (including indirect jobs) of 15.8% of the total employment in the country or 2,655,500 jobs.

According to [5], the key success factors in Mediterranean tourism designated the critical competitive variables: innovation and the willingness to change, the introduction of organizational innovations, the technological modernization of processes, facilities, and products, and the improvement of information technology. In many cases, information technology and innovation are combined to give specific markets or players competitive advantages. Apart from that, tourism loyalty is one of the most important metrics for all touristic destinations around the globe. Although the tourist industry has focused its efforts mainly on attracting new tourists, it is also true that loyalty to a tourist destination is a core goal for firms in the sector and a key source of competitive advantage [6]. Loyal tourists are likely to spend more time at a destination, better promote it and consume more goods, although the relationship may also be non-linear [7]. In this paper, authors present a platform devoted to support the whole life-cycle of tourism loyalty after the first visit. The system is designed to collect data from the initial visit by means of pervasive approaches, delivers accurate after visit data and, once returned, offers relevant recommendations based on positioning and bio-inspired recommender systems. Administrated by local tourism organizations, it adopts the customer relationship management (CRM) approach to improve tourism loyalty and overall performance. This paper provides a description of the system and is structured as follows. Section 2 presents related work on the three pillars of the system, namely CRM, recommender systems and semantic technologies. In Section 3 authors introduce the system including its architecture and main features. Section 4 demonstrates the validation of the solution performed in a controlled environment. Finally, Section 5 wraps up the paper and presents main conclusions and future work.

2. Related works

In this section, authors focus on the three principal axes of the solution. The first pillar is CRM and its particularization to tourism milieus, the second is decision support systems and their influence in tourism environments and the third is semantic technologies and their use in opinion mining scenarios. In what follows, these fields of study are analysed.

2.1. Customer relationship management in tourism

The term “customer relationship management” emerged in the information technology vendor and practitioner communities in the mid-1990s [8]. From a technological perspective, CRM systems are information systems that enable organizations to contact customers, provide services for them, collect and store customer information and analyse that information to provide a comprehensive view of the customers [9]. In other words, CRM technology applications link front office (e.g. sales, marketing and customer service) and back office (e.g. financial, operations, logistics and human resources) functions with the company’s customer “touch points” [10]. As reported in the literature (e.g. [11]), and following the technological perspective previously adopted, the market for CRM software is becoming stable almost all solutions present two kinds of deployments: On-premise and cloud-based.

Regarding to the applications of CRM in sectors, literature reports efforts in functional fields like banking (e.g. [12,13]), healthcare (e.g. [14,15]) or e-commerce (e.g. [16]), naming just some of the many instantiations of the phenomenon. Focussing on tourism, CRM is also an important factor for the tourism industry. Not in vain, according to [17], travel and tourism industry is one of the most progressive industries in information technology exploitation, providing an ideal context to investigate the influence of sophisticated CRM systems in marketing. Sigala [18] underlines two reasons behind the importance of CRM for the industry. Firstly, travellers become more price sensitive, less brand loyal and more sophisticated in their demands and tastes. Secondly, the tourism industry is experiencing globalization, competition, rising customer purchase costs, as well as increasing customers’ expectations. As a consequence of this, literature also reported a panoply of cases on the use of CRM in all tourism industry and services. Focusing, for instance, on the hotel industry there exist studies on aspects like critical success factors for CRM implementation in hotels [19], balance between relationship and revenue in hotels with regard to CRM implementations [20], relationships between CRM and business performance in hotels [21,22], functional integration and CRM implementation studies [23] and electronic CRM and service quality connections [24], citing just some of the most recent and important works. In general, literature supports the use of CRM in tourism industry but gives special importance to the need to develop recommender systems based on data gathered in order to provide relevant and personalized recommendations to both tourists and service providers alike.

2.2. Recommender systems for tourism

Recommender systems have been extensively utilized as a way of reducing the information overload and offering travel recommendations to tourists [25]. A general definition of recommender systems is the one provided in [26], defining

recommender systems as programs aimed to recommend the most suitable items (products or services) to particular users (individuals or businesses) by predicting a user's interest in an item based on related information about the items, the users and the interactions between items and users. Back in 2005, Adomavicius and Tuzhilin provided a survey of recommender systems as well as described various limitations of current recommendation methods, and discussed possible extensions that can improve recommendation capabilities and make recommender systems applicable to an even broader range of applications [27]. Around a decade after this crucial work, in [26], authors underline the importance of social information and social filtering, followers, followed, trust, reputation, credibility, content-based filtering of social data; social tagging and taxonomies in the development of recommender systems. As a consequence of this and other developments, new advanced recommendation approaches have been proposed, such as social network-based recommender systems, fuzzy recommender systems, context awareness-based recommender systems and group recommender systems; a recent and up-to-date review can be found in the works of Lu et al. [28].

Due to the importance of the topic and the increasing applicability due to the expansion of internet penetration and use, recommender systems appear in various fields of application including movie recommendation studies [29], e-business recommendations [30], expert recommendation [31,32], smart homes management [33], financial recommendations [34,35], social networks [36] or consumer electronics [37] citing just some of the most relevant cases.

Within tourism domain literature has reported many applications of recommender systems in the field. Focusing just on the most relevant and recent works, authors can cite iTravel, recommender system in mobile peer-to-peer environment [38], consensus-driven group recommender system [39], touristic tour recommendation [40], social-based touristic recommender system [41] or e-Government recommender system for tourism [42], citing just some of the most important works on the topic.

A recent survey on recommender systems on tourism [43] underlines the increasing importance of mobile interfaces for modern recommenders, although in pre-visit approaches, normally web-based ones are still prevalent. This asseveration is also supported by recent surveys devoted to mobile recommender systems [25]. Going back to the work on recommender systems in tourism [43], authors also underline that tourism recommender systems are integrating images, pictures and interactive maps. However, the strongest conclusion provided in the work might be the one underlining that the recommendation process is a crucial aspect in tourism advisory systems.

2.3. Semantic technologies and their use in opinion mining

The semantic web is active in an increasing number of functional scenarios. It enables the Web as a collaborative space for storing and querying structured data in a decentralized way where everyone can access and contribute [44]. As a consequence of its development, the role of semantic technologies in the collaborative web has become more relevant and widespread [45], particularly in the field of linked data [46]. Ontologies are the fundamental artefact behind the development of semantic technologies. According to [47], an ontology is "a formal and explicit specification of a shared conceptualisation". Ontologies formalize five kinds of components namely classes, relations, attributes, axioms and instances [48]. According to [49], among other applications, the use of ontologies may help overcome the limitations of traditional methods in natural language processing, including opinion mining technologies. Successful stories on its application in the opinion mining scenario can be found in the recent literature [50–53] on the topic.

This paper presents a social and context-aware mobile recommender system for tourists that performs its recommendations based on data from previous visits while combining also aspects taken from the social environment of the tourist. The system is built to provide efficient pattern detection and fast adaptability to changes including new Points of Interests (POI), hypes but also relevant recommendation from Destination Management Organizations (DMO). Moreover the system, apart from providing relevant recommendations to tourists, enables the relationships, by means of sound CRM systems, between DMOs and tourists in their several stages: pre-visit, during-visit and post-visit. To do so, the system, POST-VIA 360, combines semantic technologies, opinion mining, pervasive environments and artificial immune systems as the cornerstones of its technological background.

3. POST-VIA 360: system description

POST-VIA 360 is built to assist tourists in all stages of their travels: pre-visit, during-visit and post-visit by means of a CRM approach. But it is also aimed to improve quality in the all value chain by means of a continuous improvement mechanism aimed to inform DMOs and owners about relevant events, opinions or simply behaviours associated with POIs under their supervision. Built as an evolution of previous projects [54], the novelty of this approach is the inclusion of an enhanced recommender system built upon artificial immune systems, the inclusion of geo-representation enabled by PostGIS and the enhancement of opinion mining by means of feature-based opinion mining through ontologies [55] and opinion summarization [56].

Fig. 1 depicts a three-layer scheme that represents POST-VIA 360 architecture along with its main interactions and information flows. The architecture follows best practices in distributed systems design in order to assure dependability and proper service levels [57]. The three main actors of the system are: DMOs, POI administrators and tourists. The first actor deals with the system and communicates specific recommendations and tips to tourists along with overall supervision of

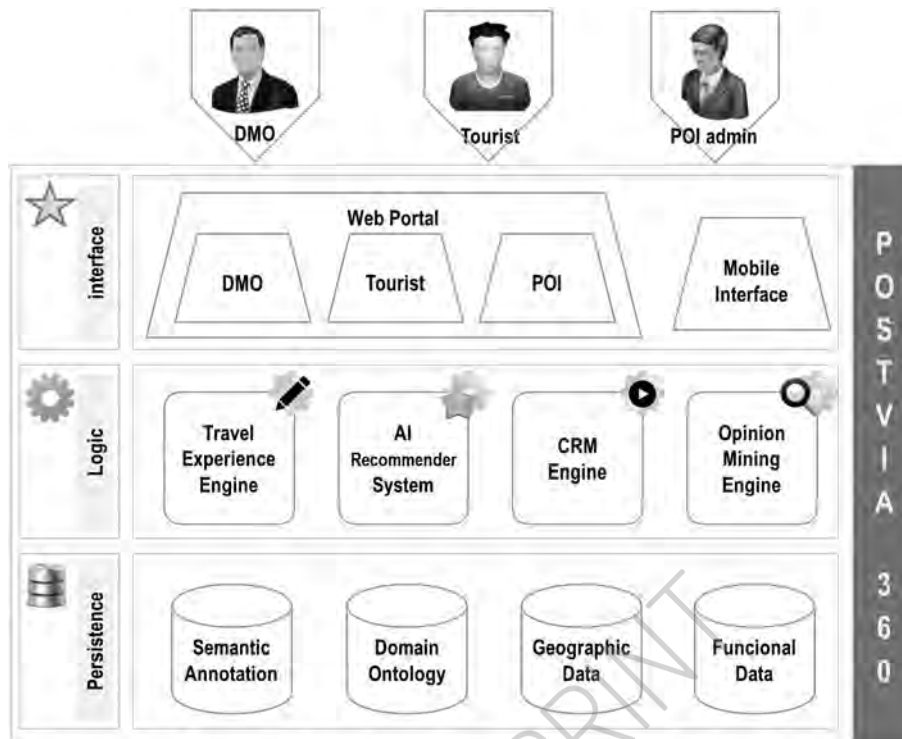


Fig. 1. POST-VIA 360 architecture, components and interactions.

the system and communication with POIs. Some of these points are under direct supervision of DMOs, like museums, for instance, while others, like restaurants are normally lead by POI owners or administrators. The second actor is the POI administrative, which is responsible of communicating news, offers, availabilities and general information about the POI. The last actor is the tourist that interacts with the system during pre-visit, during-visit and post-visit. In a nutshell, he or she browses information, receives recommendations, includes comments to the system or simply visits POIs. All this information is later used in the development of specific recommendations.

The architecture and components of the system are structured into three layers which provides the possibility to apply different replication settings at the operating system or the web service levels [58]. The purpose and functionality of the layers is as follows:

1. **Interface layer.** This layer provides a web interface and a configuration interface, given that is the component aimed to interact with tourists but also POIs administrators and DMOs. Two kinds of interfaces are available for the system. Firstly, traditional web interfaces. These interfaces are coded using HTML5 in order to easily adapt to all kinds of devices including mobile. Secondly, a native mobile interface is also available for tourists. Regarding functional subsystems within the Web interface, there is a subsystem aimed at tourists (including aspects like travel album, travel experience and social network). A second subsystem provides POIs management that can be used to maintain basic POIs information but also to access POIs metrics and other elaborated scorecard-like metrics. Finally, there is a subsystem aimed to show and maintain DMO information including managed POIs, specific DMO information, scorecard information about the use and rating of its management services, but also an interface to classic CRM functions.
2. **Business logic layer.** This layer encapsulates the main functionalities in POST-VIA 360 by encoding real-world business rules affecting the system. This business logic layer consists of four components as depicted in Fig. 1. These components are explained in what follows:
 - a. **Travel experience engine.** This component enables the valuation of the travel experience by the users before, during and after the travel. It also includes the possibility of creating a multimedia album associated to the trip including relevant information as tickets, brochures, weather conditions, geolocation-enabled itineraries but also opinions, check-ins and other personal information.
 - b. **CRM engine.** This component includes all traditional operational CRM features like, for instance, sales, marketing and service management. Traditional CRM approaches have been enriched by the use of ontologies. A view of the main features of the component is explained in detail in [54]. The semantic model includes several entities (user, profile, visit, valuation, concept, POI, similarity, content, promotion ...) as well as a set of relations (has_profile, has_friend, has_visited, has_user_valuation, promotes, manages ...). The use of semantic infrastructures allows querying the information using SPARQL queries via a specific interface designed for the system that receives the information

required and is in charge of generating the appropriate SPARQL query to access the semantic model. Regarding the storage authors adopted Jena API as main framework to load the ontology model and store the data.

- c. **Artificial immune recommender system.** Artificial immune systems is a field of study inspired by the human immune system that emerged in the 1990s as a new computational research area inspired by theoretical immunology and observed immune functions, principles and models [59]. The use of these kind of systems have been reported by literature [60] with promising results compared to other equivalent solutions. This is the reason behind the use of such systems in POST-VIA360. In this case, given that the system will be used in a mobile setup with spatial features, it is needed to add spatial information to antigens and antibodies to be able to perform spatial queries that requires the use of areas or distances. To do so, a spatial database management system as PostGIS/PostgreSQL is adopted to manage information. Regarding the model, POIs and visits include spatial information in order to be able to include geographic data in the recommender engine. The general design of the algorithm is adapted from the works of Cabanas-Abascal et al. [60], a work that also included an adaptation of artificial immune systems to recommender systems adding spatial information to the recommendation engine in a tourism scenario. However, in order to meet the specific characteristics present in POST-VIA360, specific adjustments were made in the algorithm parameters. These parameters are depicted in Table 1. Adaptation was made taking into account the dataset, POST-VIA360 architecture and overall throughput. In a nutshell, the differences of the approach implemented in POST-VIA360 with regards to the one adopted in Cabanas-Abascal et al. [60] are based in the scope of the system and, as a consequence, in the recommendation and the dataset. Algorithmics are, as already reported, similar and their parameters are adjusted from the ones adopted by Cabanas-Abascal et al. [60]. However, the scope is broader in our case. POIs are not just hotels, like in the previous effort; they include hotels, bars, museums, sights... This makes the algorithm different in the information processed and the kind of recommendation provided. With regard to the recommendation algorithm for a POI, formula is also adapted from the work of Cabanas-Abascal et al. and it is as follows:

$$poi_i = \bar{u} + \frac{\sum_{v \in N} r_{uv} (v_i - \bar{v})}{\sum_{v \in N} r_{uv}}$$

where

- “ poi_i ”: Estimated rating for POI i
 - “ \bar{u} ”: Mean of antigen’s “ u ” ratings
 - “ r_{uv} ”: Pearson’s correlation measure between the antibody and the antigen.
 - “ v_i ”: Valuation given by antibody “ v ” to POI “ i ”
 - “ \bar{v} ”: Mean of antibody’s “ v ” ratings
 - “ N ”: Set of antibodies that take part in the recommendation.
- d. **Opinion mining engine.** Opinion mining engine is based on the use of ontologies. Given that, apart from star-based ratings, users provide opinions on POIs that can be connected or not with their ratings (that, in many times are absent), it is needed to design and deploy an engine to extract information from this data. In this case, the approach adopted is inspired by the works of Peñalver-Martinez et al. [55] adapted to the touristic domain and more specifically to the ontology defined and presented in [54]. Based in the framework presented by Peñalver-Martinez [55], the engine developed for the system integrates the following modules:
- i. Natural language processing (NLP) module. This module is aimed to obtain the morphological and syntactical structure of sentences. It includes a set of tools and components developed on the Stanford Log-linear Part-Of-Speech Tagger framework.
 - ii. Semantic-based feature identification module. The module is aimed to perform a double task. Firstly, and using a domain ontology, features included in the opinions later transformed in the corpus are extracted from it. Secondly, a score that represents the importance of a given feature is calculated. This latter task is achieved by grouping features in accordance with their semantic distance and then attached to a main concept of the ontology.
 - iii. Polarity identification module. In order to identify positive, negative and neutral values of verbs, nouns and adjectives, SentiWordNet 3.0 is integrated in the polarity identification module. Final polarity is calculated by means of an intelligent average of the positive and negative senses of the set of words in the SentiWordNet database.
 - iv. Opinion mining module. This module is designed to provide effective feature sentiment classification. Features are represented by means of their Euclidean vectors and polarity is calculated taking into account positive, negative and neutral directions coded in the vector.

This opinion mining engine is used to enrich POI ratings and feed the recommendation mechanism with their enriched features.

3. **Persistence layer.** This layer stores knowledge of tourists, POIs and DMOs, with the relations among them in terms of visits, ratings and comments. Most of the characteristics of the POIs and tourists are defined by means of OWL ontologies. Based on ontologies and knowledge representation techniques, this component provides meaning to POST-VIA 360 concepts. This layer also stores relations about the different concepts which make up the platform. Some of the features are also coded using spatial databases (PostGIS) and some others are simply coded in a relational database.

Table 1
Artificial immune system parameter configuration.

Parameter	Value
Distance to nearest ones	200
# of similar elements considered in recommendation	5
Minimum population size	2500
Antibody's life decreasing value	1
Distance in mutation mobility	5
Change grade in valuation mutation	2
Antibody's life	50
Valuations per antibody	25
Default value for nooverlapdefault in Pearson measure	0
Default value for nooverlapdefault in ZeroVarianceDefault	0
Default value for overlap penalization	No penalty
Recommendation value when it can be generated	Antigen's valuations average
Number of antigens	300
Number of cycles	100

4. Validation

In this section, authors describe the validation of the tool. The final aim of this study is to determine whether POST-VIA 360 serves as a valid recommendation system in a controlled environment.

4.1. Design

Once POST-VIA 360 had been developed and tested, the second step was to test the validity of the tool in a real scenario to measure the accuracy of its recommendations. Given that the overall aim is to test the accuracy of the recommender system, a traditional approach to test these systems was adopted.

The evaluation was conducted in two steps. In the first step, the system is used by a set of users during a trip to a city different from the one they live in. A set of data was also collected from users regarding their preferences by means of a questionnaire. The aim of this first step was to provide a set of test cases in order to execute the second step. To do so, a set of users were asked to fill in the questionnaires and use the system for a weekend in a short weekend trip to Salamanca, Spain (asking the system for at least seven recommendations for each user regarding restaurants, cafes and bars). In the second step, and, by means of a GPS manipulation program, a set of recommendations for restaurants and bars will be performed in the area of Madrid city centre (Spain). POST-VIA 360 will provide a ranked set of recommendations for these users. At the same time, a set of experts also provided a ranked set of recommendation for these users in order to compare their results with the ones provided by POST-VIA 360. The expert recommendations would be carried out by a group of experts using the Delphi method—a tool that is able to allow agreement between the experts' opinions. Thus, two kinds of recommendations are produced in the validation phase. The first is provided by POST-VIA 360 and the second by the experts. Each of them recommends three different POIs for every situation given. A situation entails a geographical point in the map where the user is supposed to be situated, a given time and a track of previous actions by this user during this day (minimum three). Three different situations are provided by researchers as a testbed. POIs were imported including bars and restaurants in Madrid city centre and are aimed to be within walking distance.

To wrap-up, every situation and user will have six different recommendations, three provided by the experts and remaining three that are provided by the system. Next, in order to determine the accuracy of the results provided by the system, Precision and Recall and F1 measures were used to measure the degree of relevance of the recommendations provided by POST-VIA 360 with respect to the experts' recommendations. These measures must be complemented in order to get a full view of the recommender system from an evaluation perspective. Thus, authors decided also to perform a coverage test. The coverage of a system is a measure of the domain of items in the system over which the system can form predictions or make recommendations (i.e., the percentage of items for which a recommender system can provide recommendations) [61].

4.2. Sample

The sample is divided into two different subsets: system's users and experts. In the first set, 14 individuals took part in the experiment (8 men and 6 women). The average age of users was 29.6 years. Subjects were selected from those who answered positively to a personal invitation sent by the authors to personal contacts aimed to take a trip to Salamanca and to participate in the study. This first set of participants performed the task in a time frame of 3 months.

Table 2
Recommendations: POST-VIA 360 vs. experts.

Situation	User	Experts			POST-VIA 360		
		#1	#2	#3	#1	#2	#3
1	1	35	38	12	38	39	35
1	2	50	51	35	52	50	53
1	3	5	6	11	6	5	7
1	4	68	71	69	69	70	71
1	5	38	39	12	71	73	72
1	6	75	76	79	76	75	51
1	7	17	20	38	25	21	17
1	8	8	6	11	6	5	7
1	9	76	75	7	9	7	8
1	10	34	49	46	39	35	37
1	11	34	49	46	39	35	37
1	12	67	76	73	70	75	76
2	1	49	47	31	49	45	46
2	2	61	63	19	61	65	63
2	3	8	9	13	9	8	4
2	4	74	72	71	73	72	74
2	5	74	72	71	73	72	74
2	6	76	75	79	75	76	52
2	7	18	22	37	22	23	18
2	8	8	9	13	9	8	4
2	9	10	3	2	1	4	3
2	10	48	49	31	49	47	48
2	11	48	45	31	49	47	48
2	12	74	72	71	69	72	68
3	1	84	87	90	90	88	89
3	2	96	97	86	97	95	98
3	3	98	99	95	99	96	98
3	4	90	87	88	89	87	90
3	5	90	88	87	88	89	90
3	6	88	90	87	90	88	89
3	7	98	99	95	99	96	98
3	8	98	99	95	99	96	98
3	9	98	99	95	99	96	98
3	10	96	97	86	97	95	98
3	11	96	97	86	97	95	98
3	12	90	87	88	89	87	90

Secondly, to provide recommendations, a group of four experts was invited to participate in the study. They all serve as professional tourism guides in the city of Madrid. All of them were invited to participate by researchers who contacted them using their personal networks.

4.3. Results and discussion

Once the information from users was gathered by the system and the questionnaire was sent to experts, along with a briefing on the actions performed by the users, experts were gathered to submit their recommendations.

This classification was carried out by the four experts using the Delphi method. Each expert issued a pair of recommendations for every user and situation. Kendall's coefficient of concordance (W) was computed to measure the level of agreement among the experts. The four experts obtained a coefficient of agreement of $W = 0.5938$. Finally, agreement between experts was achieved by applying the Delphi method, producing an output of three recommendations per situation and user. This is a moderate agreement, however, due to the high number of POIs and the important number of recommendations, this could be acceptable.

A set of recommendations were also obtained from POST-VIA 360. This set is composed by three different recommendations per situation (three different situations) and user (twelve users), reaching the number of 108 recommendations made by the system and an equal number performed by experts. Table 2 shows results of recommendations process performed by both the POST-VIA 360 tool and the experts.

Table 2 presents data for the 3 situations described. For each situation, data for the 12 experts is presented. In sum, 36 situations have been tested. For each situation, 3 recommendations were obtained from experts and 3 more from POST-VIA360. In general, there is a remarkable similarity in the recommendations comparing experts and POST-VIA360 for all situation and expert pairs. However, to evaluate the quality of recommendations from the two sources in a scientific way, precision, recall and F1 were used as measures. These measures are widely used in the literature to evaluate the quality of recommender systems. The recall-metric indicates the effectiveness of a method for locating interesting topics. In the case of POST-VIA 360 recommendations are POIs. The precision-metric represents the extent to which the items recommended

Table 3
Comparison of POST-VIA 360's accuracy with that of similar works.

Reference	Precision	Recall	F1
POST-VIA 360 (first option only)	0.056	0.056	0.056
POST-VIA 360 (whole set)	0.917	0.306	0.458
[60]	0.864	1.000	0.927
[62]	0.480	0.480	0.480
[63]	0.880	0.594	0.698
[64]	0.714	0.771	0.741

by the system are interesting to users compared to the ones performed by experts. F1 is calculated as the sum of the weights of precision and recall.

In the first analysis performed, the recommendations by the experts were considered correct. So, if the product recommended by POST-VIA 360 coincides with that recommended by the experts (the first recommendation on every set) then it was considered a correct recommendation. Firstly, precision score determines the number of correct recommendations offered by the system with respect to the overall recommendations offered by POST-VIA 360. Recall score represents the relation between the number of correct recommendations with respect to the overall correct products (in this first analysis the products selected by the expert). In the case studied here, precision, and recall and, as a consequence, F1 values are exactly the same because the number of correct recommendations is equal to the number of recommendations offered by the system. In this first analysis, these values reach 0.056, a very modest figure in terms of accuracy. The reason behind this low accuracy could root on the big amount of POIs available for a recommendation and the relative similarity among them. The first argument that supports this low accuracy is the high number of POIs available for recommendation. The sample included a set of 100 POIs. This is a considerable amount of data to choose from without previous information on these specific POIs and their interaction with users. Given that POST-VIA360 is aimed to work in a setting in which recommendations will be made on the basis of previous visits to a specific place, authors are confident in increasing the accuracy in these situations. However, authors believe that the most important argument supporting the modest accuracy is the composition of the sample. While there are around 10 different categories on the set (pubs, thematic restaurant, gourmet restaurant, buffet, hype restaurant, bar...) the real problem is caused by the similarity of POIs inside these categories. In a nutshell, it is important to note that the differences between POIs in the same category, price and area are just a matter of taste for experts and tourists. Given that Madrid is a city with a huge offer in the culinary arena and that, for instance in the traditional restaurant category, there are many similar restaurants with similar prices and placed in the same area, it is logical to have recommendations in all of them in similar situations. This is a circumstance that lead to the modest accuracy figures obtained in the test.

This fact leads us to the second analysis performed. In this analysis, the two sets of recommendations are compared as a whole. In this case, a whole set of recommendations is considered valid if one of its 3 values is present in the recommendation provided by experts. In this case, precision reaches a remarkable 0.917 while recall achieves 0.306 and, finally, F1 reaches 0.458. In this case, the excellent results in terms of precision are moderated by the humble figures on the recall side.

As a second source of validation figures, authors consider of interest the measurement of the usefulness of recommendations. It is obtained by means of the coverage, a measure of the domain of items in the system over which the system can form predictions or make recommendations or, in other words, the percentage of a dataset that the recommender system is able to provide predictions for. Taking into account the set of recommendations made by POST-VIA 360, its coverage reaches 48%.

Regarding the comparison with relevant literature there is some research in diverse functional areas are available for a comparison, although there are limited examples in the functional field of tourism. Table 3 shows the comparison of POST-VIA 360's results with recent efforts in the literature regarding recommendations in the tourism scenario.

Results in Table 3 picture POST-VIA 360 in the same trend as in previous efforts, with some aspects even above previous figures and some others below their results. Authors listed four works published in the field of recommender systems for tourism that provided precision, recall and F1 measures. The sizes of databases and samples are not comparable (although, for instance [62] is quite similar), however, the domain and the experiments can be compared and provide accurate information to be compared with. Taking this information into account, precision reaches 0.917—the highest of the compared systems, while recall is reaching a modest 0.306 and, as a consequence, F1 measure is moderated, although aligned with previous efforts, e.g., [62]. However, the reason behind this humble recall figures is double. Firstly, the considerable size of the POIs sample (that is not limited in terms of type of POI for a given recommendation, like in some of the efforts in similar works) and secondly, the fact that recall values are calculated taking into account specific POIs and not categories of them that, in our case, can be restaurants, bars and pubs, for instance. With regards to the remarkable precision results obtained in the whole set, the reason behind these results is twofold. Firstly, the evident decision to take the three recommendations as a whole. This makes better precision figures possible while simultaneously leading to lower levels of recall. Secondly, it is also important to note that, although the similarity of the experiments is moderate, artificial immune algorithm is providing good (but similar) results in the scenario provided compared to other approaches adopted in previous works.

In spite of the encouraging results, there is still room for improvement. A closer look at results (with the help of some experts) shows that experts provided suggested recommendations also for POIs that are not representing a clear category,

while the system is ignoring these POIs for its recommendations. Although a more in-depth analysis regarding the lack of accuracy in recommendations must be performed, it seems that a better knowledge representation must be provided or better data will have to be provided to the system. Given that, by nature, POST-VIA 360 is maintained by DMOs, the reason behind this lack of coverage should be routed in the need to have up-to-date data on the side of DMOs to feed the system. In this case, experts seem to have information that is not provided to the system.

In this study, internal or external validity threats are present. Regarding internal validity, it is concerned with correctly concluding that an independent variable is, in fact, responsible for variation in the dependent variable, in this case, recommendations. Experts may not have a comparable level of knowledge or expertise. Given that experts were selected because of their expertise, experience and current work, authors expect that the respondents have a comparable level of knowledge and expertise. However, to ensure a similar level of knowledge among experts, aspects like professional certification should be taken into account. Conversely, this aspect is beyond the scope of this work.

External validity is concerned with the generalizability of research findings to and across populations of participants and settings. The authors face two possible threats. The first is the limited number of subjects, which complicates generalization of the results. The second is subject representativeness, given that the sample was not taken randomly. Future works will tackle both threats.

5. Conclusions and outlook

The tourism industry is undergoing a revamp, advanced by the use of technologies by consumers and service providers alike. Tourists have become the paradigm of prosumers (consumers + producers, tourists like to voice their opinions to businesses hence helping them improve on the level of their service) and their opinions and inputs are used to improve services but also to influence other tourists in their decisions. This phenomenon is flooding the internet with a huge amount of data available for tourists and service providers alike. Moreover, current mobile data explosion but also recent forecasts, including Cisco's annual Global Mobile Data Traffic Forecast published in February 2015, unveil an increasing volume of data coming from these devices. In this scenario, information overload is a threat for all actors in tourism that recommender systems must analyse in order to provide only highly-relevant information. In this paper, authors presented POST-VIA 360, a system designed to analyse all information coming from the whole life-cycle of tourism. POST-VIA 360 is aimed to provide support for tourists, but also POIs managers and DMOs alike. Built upon cutting edge technologies, including artificial immune systems, geo-representation, feature-based opinion mining, semantic technologies and opinion summarization naming just of the most important aspects, POST-VIA 360 has been validated in real contexts and has shown remarkable results in touristic recommendations.

In spite of its initial benefits, POST-VIA 360 presents some limitations. The first and maybe the most important is the scope of data the tool is able to analyse. Nowadays social media users are producing huge amount of information in the social web using Facebook, Twitter, Instagram or LinkedIn, but these sources of information are currently out of reach for the system. From the technical point of view POST-VIA 360 is dependent on a given dataset and, to some extent, on a given domain. Finally, regarding the testing made, authors assume that the introductory study provided presents some threats as highlighted previously in this paper.

Future research will be threefold. Firstly, it is aimed to expand the possibilities of the AI recommender system by adjusting the algorithm in order to improve its efficiency and by adding mechanism to ensure that POST-VIA 360 will be independent from a given dataset. Secondly, it is aimed to expand the possibilities of text summarization techniques by the inclusion of specific micro-blog natural language interpreters and Facebook status updates to boost the performance in these social networks. Finally, authors would also investigate the use of specific Social CRM tools for DMOs to promote POIs under their supervision designed to automatically detect social web hypes and link them to events and campaigns launched over specific POIs.

Acknowledgment

This work is supported by the Spanish Ministry of Industry, Tourism, and Commerce under the project "*Diseño, desarrollo y prototipado de una plataforma TIC de servicios post-viaje a los turistas*" (IPT-2011-0973-410000).

References

- [1] United Nations World Tourism Organization, UNWTO tourism highlights, 2014 edition, Madrid, 2014.
- [2] Eurostat, Tourism statistics, 2015.
- [3] European Commission, Flash Eurobarometer 414: Preferences of Europeans towards tourism 2015, Vol. 414, 2015.
- [4] World Travel and Tourism Council, Travel & Tourism. Economic Impact 2014 Spain, London, UK, 2014.
- [5] C. Camisón, V.M. Monfort-Mir, Measuring innovation in tourism from the Schumpeterian and the dynamic-capabilities perspectives, *Tourism Manag.* 33 (4) (2012) 776–789.
- [6] C. Antón, C. Camarero, M. Laguna-García, Towards a new approach of destination loyalty drivers: satisfaction, visit intensity and tourist motivations, *Curr. Issues Tour.* 0 (0) (2014) 1–23.
- [7] M. Meleddu, R. Paci, M. Pulina, Repeated behaviour and destination loyalty, *Tourism Manag.* 50 (2015) 159–171.
- [8] A. Payne, P. Frow, A strategic framework for customer relationship management, *J. Mark.* 69 (4) (2005) 167–176.
- [9] F. Khodakarami, Y.E. Chan, Exploring the role of customer relationship management (CRM) systems in customer knowledge creation, *Inf. Manage.* 51 (1) (2014) 27–42.

- [10] L. Fickel, Know your customer, *CIO Mag.* 12 (21) (1999) 62–72.
- [11] R. Colomo-Palacios, J.M.Á. Rodríguez, Semantic representation and computation of cloud-based customer relationship management solutions, in: R. Meersman, H. Panetto, A. Mishra, R. Valencia-García, A.L. Soares, I. Ciuciu, F. Ferri, G. Weichhart, T. Moser, M. Bezzi, H. Chan (Eds.), *On the Move to Meaningful Internet Systems: OTM 2014 Workshops*, Springer, Berlin, Heidelberg, 2014, pp. 347–357.
- [12] A. Krasnikov, S. Jayachandran, V. Kumar, The Impact of Customer Relationship Management Implementation on Cost and Profit Efficiencies: Evidence from the US Commercial Banking Industry, *J. Mark.* 73 (6) (2009) 61–76.
- [13] F.N. Ogwueleka, S. Misra, R. Colomo-Palacios, L. Fernandez, Neural network and classification approach in identifying customer behavior in the banking sector: A case study of an international bank, *Hum. Factors Ergon. Manuf. Serv. Ind.* 25 (1) (2015) 28–42.
- [14] M.-L. Wang, An evaluation of customer relationship management in hospital-based and privately run nursing homes in Taiwan, *Total Qual. Manage. Bus. Excell.* 24 (9–10) (2013) 1004–1021.
- [15] Subash Lonial, P.S. Raju, Impact of service attributes on customer satisfaction and loyalty in a healthcare context, *Leadersh. Health Serv.* 28 (2) (2015) 149–166.
- [16] Á. García-Crespo, R. Colomo-Palacios, J.M. Gómez-Berbis, B. Ruiz-Mezcua, SEMO: a framework for customer social networks analysis based on semantics, *J. Inf. Technol.* 25 (2) (2010) 178–188.
- [17] G. Mastorakis, N. Trihas, E. Perakakis, I. Kopanakis, E-CRM in tourism exploiting emerging information and communication technologies, *Anatolia* 26 (1) (2015) 32–44.
- [18] M. Sigala, Integrating customer relationship management in hotel operations: managerial and operational implications, *Int. J. Hosp. Manag.* 24 (3) (2005) 391–413.
- [19] A. Padilla-Meléndez, A. Garrido-Moreno, Customer relationship management in hotels: examining critical success factors, *Curr. Issues Tour.* 17 (5) (2014) 387–396.
- [20] X.L. Wang, Relationship or revenue: Potential management conflicts between customer relationship management and hotel revenue management, *Int. J. Hosp. Manag.* 31 (3) (2012) 864–874.
- [21] S.-I. Wu, C.-L. Lu, The relationship between CRM, RM, and business performance: A study of the hotel industry in Taiwan, *Int. J. Hosp. Manag.* 31 (1) (2012) 276–285.
- [22] A. Josiassen, A.G. Assaf, L.K. Cvelbar, CRM and the bottom line: Do all CRM dimensions affect firm performance? *Int. J. Hosp. Manag.* 36 (2014) 130–136.
- [23] E.C.S. Ku, Functional integration and systems implementation of customer relationship management in hotel industry: a multilevel analysis, *Int. J. Inf. Technol. Decis. Mak.* 13 (01) (2013) 175–196.
- [24] J. Tian, S. Wang, Signaling Service Quality via Website e-CRM Features More Gains for Smaller and Lesser Known Hotels, *J. Hosp. Tour. Res.* (2014) 1096348014525634.
- [25] D. Gavalas, C. Konstantopoulos, K. Mastakas, G. Pantziou, Mobile recommender systems in tourism, *J. Netw. Comput. Appl.* 39 (2014) 319–333.
- [26] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, Recommender systems survey, *Knowl.-Based Syst.* 46 (2013) 109–132.
- [27] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, *IEEE Trans. Knowl. Data Eng.* 17 (6) (2005) 734–749.
- [28] J. Lu, D. Wu, M. Mao, W. Wang, G. Zhang, Recommender system application developments: A survey, *Decis. Support Syst.* 74 (2015) 12–32.
- [29] L.O. Colombo-Mendoza, R. Valencia-García, A. Rodríguez-González, G. Alor-Hernández, J.J. Samper-Zapater, *RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes*, *Expert Syst. Appl.* 42 (3) (2015) 1202–1222.
- [30] D. Wu, G. Zhang, J. Lu, A fuzzy preference tree-based recommender system for personalized business-to-business E-Services, *IEEE Trans. Fuzzy Syst.* 23 (1) (2015) 29–43.
- [31] R. Valencia-García, F. García-Sánchez, C. Casado-Lumbreras, D. Castellanos-Nieves, J.T. Fernández-Breis, Informal learning through expertise mining in the social web, *Behav. Inf. Technol.* 31 (8) (2012) 757–766.
- [32] R. Colomo-Palacios, C. Casado-Lumbreras, P. Soto-Acosta, S. Misra, Providing knowledge recommendations: an approach for informal electronic mentoring, *Interact. Learn. Environ.* 22 (2) (2014) 221–240.
- [33] K. Rasch, An unsupervised recommender system for smart homes, *J. Ambient Intell. Smart Environ.* 6 (1) (2014) 21–37.
- [34] A. Rodríguez-González, R. Colomo-Palacios, F. Guldres-Iglesias, J.M. Gómez-Berbis, A. García-Crespo, FAST: Fundamental analysis support for financial statements. Using semantics for trading recommendations, *Inf. Syst. Front.* 14 (5) (2012) 999–1017.
- [35] I. Gonzalez-Carrasco, R. Colomo-Palacios, J.L. Lopez-Cuadrado, A. García-Crespo, B. Ruiz-Mezcua, PB-ADVISO: A private banking multi-investment portfolio advisor, *Inform. Sci.* 206 (2012) 63–82.
- [36] V. Stantchev, L. Prieto-González, G. Tamm, Cloud computing service for knowledge assessment and studies recommendation in crowdsourcing and collaborative learning environments based on social network analysis, *Comput. Hum. Behav.* 51 (Part B) (2015) 762–770.
- [37] B. Barragáns-Martínez, E. Costa-Montenegro, J. Juncal-Martínez, Developing a recommender system in a consumer electronic device, *Expert Syst. Appl.* 42 (9) (2015) 4216–4228.
- [38] W.-S. Yang, S.-Y. Hwang, iTravel: A recommender system in mobile peer-to-peer environment, *J. Syst. Softw.* 86 (1) (2013) 12–20.
- [39] J. Castro, F.J. Quesada, I. Palomares, L. Martínez, A consensus-driven group recommender system, *Int. J. Intell. Syst.* 30 (8) (2015) 887–906.
- [40] I. García-Magariño, ABSTUR: An agent-based simulator for tourist urban routes, *Expert Syst. Appl.* 42 (12) (2015) 5287–5302.
- [41] A. García-Crespo, J. Chamizo, I. Rivera, M. Mencke, R. Colomo-Palacios, J.M. Gómez-Berbis, SPETA: Social pervasive e-Tourism advisor, *Telemat. Inform.* 26 (3) (2009) 306–315.
- [42] M. Al-Hassan, H. Lu, J. Lu, A semantic enhanced hybrid recommendation approach: A case study of e-Government tourism service recommendation system, *Decis. Support Syst.* 72 (2015) 97–109.
- [43] J. Borrás, A. Moreno, A. Valls, Intelligent tourism recommender systems: A survey, *Expert Syst. Appl.* 41 (16) (2014) 7370–7389.
- [44] A. Dessi, M. Atzori, A machine-learning approach to ranking RDF properties, *Future Gener. Comput. Syst.* 54 (2016) 366–377.
- [45] R. Bonacin, N. Dessi, M.G. Fugini, O. Nabuco, M.R. da Silveira, Special issue on semantic technologies for collaborative web, *Future Gener. Comput. Syst.* 54 (2016) 344–347.
- [46] R. Colomo-Palacios, V. Stantchev, A. Rodríguez-González, Special issue on exploiting semantic technologies with particularization on linked data over grid and cloud architectures, *Future Gener. Comput. Syst.* 32 (2014) 260–262.
- [47] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: Principles and methods, *Data Knowl. Eng.* 25 (1–2) (1998) 161–197.
- [48] T. Gruber, Collective knowledge systems: Where the social web meets the semantic web, *Web Semant. Sci. Serv. Agents World Wide Web* 6 (1) (2008) 4–13.
- [49] Y. Hernández-González, C. García-Moreno, M.Á. Rodríguez-García, R. Valencia-García, F. García-Sánchez, A semantic-based platform for R&D project funding management, *Comput. Ind. Eng.* 65 (5) (2014) 850–861.
- [50] G. Bello-Orgaz, J.J. Jung, D. Camacho, Social big data: Recent achievements and new challenges, *Inf. Fusion* 28 (2016) 45–59.
- [51] E. Cambria, B. Schuller, Y. Xia, C. Havasi, New avenues in opinion mining and sentiment analysis, *IEEE Intell. Syst.* 28 (2) (2013) 15–21.
- [52] J.M. Alvarez-Rodríguez, M. Vafopoulos, J. Llorens, Enabling policy making processes by unifying and reconciling corporate names in public procurement data. The CORFU technique, *Comput. Stand. Interfaces* 41 (2015) 28–38.
- [53] M. Bohloul, J. Dalt, M. Dornhöfer, J. Zenkert, M. Fathi, Knowledge discovery from social media using big data-provided sentiment analysis (SoMABIT), *J. Inf. Sci.* 41 (6) (2015) 779–798.
- [54] R. Colomo-Palacios, A. Rodríguez-González, A. Cabanas-Abascal, J. Fernández-González, Post-via: After visit tourist services enabled by semantics, in: P. Herrero, H. Panetto, R. Meersman, T. Dillon (Eds.), *On the Move to Meaningful Internet Systems: OTM 2012 Workshops*, Springer, Berlin, Heidelberg, 2012, pp. 183–193.
- [55] I. Peñalver-Martínez, F. García-Sánchez, R. Valencia-García, M.Á. Rodríguez-García, V. Moreno, A. Fraga, J.L. Sánchez-Cervantes, Feature-based opinion mining through ontologies, *Expert Syst. Appl.* 41 (13) (2014) 5995–6008.

- [56] E. Lloret, E. Boldrini, T. Vodolazova, P. Martínez-Barco, R. Muñoz, M. Palomar, A novel concept-level approach for ultra-concise opinion summarization, *Expert Syst. Appl.* 42 (20) (2015) 7148–7156.
- [57] V. Stantchev, M. Malek, Translucent Replication for Service Level Assurance, in: *High Assurance Services Computing*, Berlin, New York, 2009, pp. 1–18.
- [58] V. Stantchev, M. Malek, Addressing Web Service Performance by Replication at the Operating System Level, in: *ICIW '08: Proceedings of the 2008 Third International Conference on Internet and Web Applications and Services*, Los Alamitos, CA, USA, 2008, pp. 696–701.
- [59] M.A. Chikh, M. Saidi, N. Settouti, Diagnosis of diabetes diseases using an artificial immune recognition system2 (AIRS2) with fuzzy k-nearest neighbor, *J. Med. Syst.* 36 (5) (2011) 2721–2729.
- [60] A. Cabanas-Abascal, E. García-Machicado, L. Prieto-González, A. de A Seco, An item based geo-recommender system inspired by artificial immune algorithms, *J. Univ. Comput. Sci.* 19 (13) (2013) 2013–2033.
- [61] J.L. Herlocker, J.A. Konstan, L.G. Terveen, J.T. Riedl, Evaluating collaborative filtering recommender systems, *ACM Trans. Inf. Syst.* 22 (1) (2004) 5–53.
- [62] Á. García-Crespo, J.L. López-Cuadrado, R. Colomo-Palacios, I. González-Carrasco, B. Ruiz-Mezcua, Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain, *Expert Syst. Appl.* 38 (10) (2011) 13310–13319.
- [63] L. Shi, F. Lin, T. Yang, J. Qi, W. Ma, S. Xu, Context-based Ontology-driven recommendation strategies for tourism in ubiquitous computing, *Wirel. Pers. Commun.* 76 (4) (2013) 731–745.
- [64] R. Schaller, M. Harvey, D. Elsweiler, Detecting event visits in urban areas via smartphone GPS data, in: M. de Rijke, T. Kenter, A.P. de Vries, C. Zhai, F. de Jong, K. Radinsky, K. Hofmann (Eds.), *Advances in Information Retrieval*, Springer International Publishing, 2014, pp. 681–686.

