The evolution of travel recommender systems: A comprehensive review

V. K. Muneer¹* and K. P. Mohamed Basheer²

Abstract
Travel has always been an innate desire of human beings. People wanted to travel irrespective of any hurdles like geographical barrier, age, gender, or colour with different motivations. Nowadays travel and adventure became the most trending entertainment as well. Planning a trip is a time-consuming and herculean task for inexperienced travelers. Here comes the possibility of expert opinion for scheduling a perfect travel plan. With the development of information technology and social media, there are numerous possibilities and opportunities in fetching suitable information which can turn out to set up an appropriate travel plan and hence enhance the quality of travel. The significance of a Recommender System (RS) comes in the picture which can address travel-related queries. Personalized Travel RS will add more customization and user-specific features than Automatic Travel RS. In this paper, we conducted a detailed review and chronological evolutions of various methods and techniques used in the travel and tourism sector and compared their efficiency in Recommendations.

Keywords
Artificial Intelligence, Recommender Systems, Autoencoders, PTRS, Travel and Tourism.

AMS Subject Classification
68T01.

1,2 Department of Computer Science, Sullamussalam Science College, Kerala-673639, India.
*Corresponding author: ¹vkmuneer@gmail.com; ²mbasheerkp@gmail.com

Article History: Received 24 July 2020; Accepted 16 October 2020 ©2020 MJM.

1. Introduction
At the point when humanity fired standing up on their feet, the excursion of movement began. The account of human travel has a history as old as human being itself. The journey of movement might be for food, hunting, business, connections, relocations, profession, amusement, or information. Individuals invest energy and cash for this. The mode of this travel might be solo, couples, friends, families, and method of movement can be by road, water, or air. Sharing online reviews and experiences of a user in purchase or travel became so easy because of the latest advancements in internet applications and widespread communication technologies. These public reviews can be read by other users that can be utilized for different strategy making and obtaining detailed information about tourism-related services. Many online platforms like websites, online shopping sites, hotels, booking applications, blogs, rating sites, social networking websites play a vital role as a channel for internet communication. To an extent, these reviews help people to choose different preferences and priorities which can significantly influence decision-making behaviors.

Several studies have been carried out to analyze the impact of online media on business promotion, education, entertainment, travel, and tourism worldwide. The photos and posts on Facebook strongly influence the people to decide on their travel plans. An online survey claims that 93% of travelers used to observe online reviews which play a significant impact on their booking decisions worldwide [1]. Another study says
The evolution of travel recommender systems: A comprehensive review — 1778/1785

that 60% of vacation nomads and 41% of commercial people arranging their plans with the help of social media. According to Google statistics, about 80 percent of people seek information from the web for planning their upcoming holidays. Another analysis uncovers that 52% remains positive while 48% reexamine their travel arrangements based on reviews in online platforms [2]. The term electronic-word-of-mouth (e-WOM) has been universally adopted by customers on several online platforms [3].

The paper is categorized as follows - Segment 1 brings an idea about why travel and tourism sector and the relevance of RS in this domain, Segment 2 introduces prominent RS approaches, algorithms used and their efficiency, Segment 3 discusses on studies and researches carried out on this area in previous years, Segment 4 point out the discussions and remarks from the study and Segment 5 concludes with the inferences.

2. Approaches Of Recommender System

Recommender Systems (RS) became key tools in the era of information overflow. The trajectory of RS begun with generic recommender engines for non-customized recommendations which offered approach to personalized RS and afterward extended to contextualized personalization with an influx of Computational Intelligence. The emerging concept like Big Data, high-performance computing, the boom in social networking also acting as decisive input for various data analytics and RS. Even though Non-personalized RS can be effective in some situations, the personalized recommendations are more meaningful and widely accepted.

Making suggestions to people based on their interests and preferences is one of the most significant features of the modern business model in online purchases, books, trips, and films. Netflix is the best example as more than 75% of movies viewed on Netflix are suggested by RS [4]. User-generated data from social media can be processed and used by the RS. Facebook is one of the biggest interpersonal communication media through which a lot of individuals associate with one another and express their perspectives. The ongoing development in Facebook has made ready to comprehend user’s choices and inclinations to make suggestions effectively [40]. The places of the user’s interest are found and ranked using opinion mining and recommended to the user. Here comes the role of the tourism recommender system which was introduced by Delgado and Davidson [5]. The prominent filtering mechanisms used in RS are Collaborative filtering, Content-based filtering, context-aware filtering and Hybrid filtering.

2.1 Collaborative Filtering (CF)

Collaborative filtering is a technique that filters information for different sets of data using different collaboration techniques. CF analyses historical transactions to create connections between users and items/products. Depending upon the relationship between users and products, the two most popular techniques to CF are the latent factor models (LFM) [6] and Neighbourhood models (NM). LFM directly describes users and items where NM analyses the similarities between products or users.

2.2 Content-Based recommendation system (CB)

This is a more user-specific filtering system in which the analysis of user’s profile, demographic data, user’s history, and interest is taken into consideration for recommending items to that user which must be similar to the ones that user already selected in past. The similarity of items is calculated based on their similarity concerning features [7]. CF utilizes the user-item communications and thus, alludes to people to people connection, though CB utilizes the attribute-information of users and items. Be that as it may, both individual methodologies experience restrictions. Information sparsity and cold-start problem, i.e., lack of adequate inputs from other users to process the recommendation [8].

2.3 Context-aware filtering

Another approach in which the context-based information is also gathered and processed for suggesting the recommendation. The efficiency of recommendation can considerably increase when additional contextual information like time, area, music, social facts are taken as input parameters along with traditional user-item values [9]. While considering the performance of the above RSs, there are few impediments such as overspecialization, cold start problem and sparsity problem [10] to be addressed.

2.4 Hybrid Recommendation

To overcome the above challenges, a collective mechanism, a hybrid approach has been introduced. The challenges faced in the above methods resolved to and extend and it could improve the efficiency and accuracy of recommendation. The methodology of HF is as follows: individual filtering is performed by CF and CB and combining the results. Then by inserting a CB filtering skill to Collaborative filtering and vice versa is also possible or by merging the techniques into one single model [11]. The diagrammatic representation of filtering mechanisms are shown in Figure 1.

![Figure 1. Classification of Filtering methods](image-url)
3. Recommender Systems In Literature

Vipul V et. al [12], proved how a hybrid collaborative filtering method performs better than a collaborative and content based approach. In 2017 Farhim Mansur et.al, conducted a detailed analysis of various approaches in RS considering all prominent types of recommendation systems, sub-modules in each of them. CF divided into model-based filtering and memory-based filtering. Memory-based filtering further divided into Item-based CF and User-based CF techniques. Item based collaborative filtering further categorized into three similarities such as cosine based, correlation based and adjusted cosine similarity [13]. The authors steered a comprehensive study on various methods used in hybrid filtering led to specifications of seven different types including a) weighted, b) switching, c) mixed approach, d) feature combination, e) cascade, f) feature augmentation, and g) meta-level hybridization approaches.

Kinjal Chaudhari and Ankit Thakkar [14] in 2019 examined various prominent recommendation methods in travel and related allies. The paper attempts to categorize travel recommendation systems into five major segments. They are Attraction Recommendations, Hotel Recommendations, Tourism Recommendations, Restaurant Recommendations, and others. Recommendations include photography spots, food, outfits, transportation, and weather recommendations. They suggested that all possible inputs, possibilities, and reviews should be taken from each of these segments to develop a complete travel recommendation system. Various domains in TRS shown by the Figure 2.

Authors conducted a Comparative analysis with existing seven prominent tourism-related surveys for 15 diverse criteria in travel and tourism. The survey discussed recommendation model includes mobile tourism recommender systems [15], mobile multimedia recommendation in smart communities [16], mobile recommender systems in tourism [17], intelligent tourism recommender systems [18], food recommender system [19], recommender system application developments [20], recommendation technologies in travel and tourism [21].

Shah Khusro et. al [22] in 2016 reported the problems in developing efficient recommender systems by analyzing the current trends, issues, challenges, and research opportunities in the domain. The popularity of RS has been extended in several areas not limited to business products, trips, videos, photos, editorials, news, and books. The authors tried to explain (i) common concepts and techniques related to different types of RS, (ii) to discuss a few of the major loopholes and challenges in RS development, (iii) to propose solutions, techniques and research advice that would help the designers of RS. Considering all possible measures, the shortlisted prominent issues in RS are Cold Start Problem, Synonymy, Shilling Attacks, Privacy and ten other problems. They proposed solutions for the challenges in RS such as using demographic filtering and clustering, it can minimize latency and increase performance. The ability to handle sparsity and filtering out obsolete items could increase the accuracy of suggestions.

Nilashi.M. et al., in 2015 proposed a new hybrid method by using two techniques, dimensionality reduction and related prediction methods for recommending hotels to users. The authors adopted the Adaptive Neuro-Fuzzy Inference System (ANFIS) as a supervised approach for building the models of prediction and they executed a clustering on data using the Expectation Maximum algorithm before constructing this model. Pearson Correlation Coefficient approach is selected for users and items similarity calculation in clustering and hence ANFIS could construct effective prediction models by clustering of data of users and objects [23]. Figure 3 give the diagramatic representation. Blogs and reviews are taken into consideration for ranking and recommending the hotels, Zulkefli et al., [24] suggested a new Hotel RS based on online travelogues. The work is not focused on microblogs such as twitter but focused on blogs where the user can express limitless write-ups. Work aims to suggest a suitable hotel that satisfies customer tastes and requirements. It contains three steps.

1. Evaluation of each hotel’s neighboring natural environments from online scripts.
2. Calculation of user preferences using their reviews of hotels that they have chosen earlier.
3. Articulate the similarity between the above two and suggest the top-n hotels to recommend to that user.

A comparison of user’s current preference with previous hotels is calculated by using Cosine Similarity Measure and top-n hotels will recommend to the user by the decreasing order, authors selected a couple of hotels from Agoda.com and Booking.com website as dataset to a conclusion that the blog information can increase the efficiency result based on hotel preferences and user preferences [24][26].

C. Paola et.al [25], introduced a methodology, comment classifier (CC) to recommend hotels by categorizing the online comments written by customers. CC has taken comments from TripAdvisor.com and classified into three groups. Good, Fair, and Bad. The major modules included in this model are gathering information, ontologies design, and comment classification. For this experiment, the authors analyzed 686 comments from 74 hotels on tripadvisor.com website. A recommender system was developed to reduce the number of efforts that a user needs to invest in identifying the right travel plan of his taste.

In 2017 Zhang and Morimoto [26] proposed a method that automatically decides appropriate attributes of vector space from hotel review comments which may be written in Natural Language. They introduced Latent Dirichlet Allocation (LDA) for text analysis and pull out representative topics about hotels automatically. This method has two phases, one is Hotel Vector
The evolution of travel recommender systems: A comprehensive review

Figure 2. Domains in Travel and Tourism Recommender system.

Figure 3. Block diagram of ANFIS Model

Generation, and the second one is Item-based Collaborative Recommendation. The pre-processing of natural language is executed by Stanford CoreNLP, which includes POS tagger, named entity recognizer, and sentiment analysis tools. Zhang and Morimoto analyzed these texts to find sentiment for each extracted topic for each hotel. These inputs are good enough to generate topic-sentiment pair for each sentence and produced recommendations. 2256 reviews from TripAdvisor taken as a data source for evaluation. The authors claimed that the performance of their method is better than Item-based CF and User-based CF.

Information collected from geo-tagged travel photos shared in social media can play a vital role in travel recommendations. Jongmyo Hana et al. [27], proposed a model, PLANNER (Personalized LANdmarks recommender) for analyzing the relationship between the significance of landmarks of travelers based on their trip’s spatial and temporal properties to generate clusters of landmark recommendations. They used a well-known photo-sharing social network, flickr, from which a total of 372,000 photos collected for the experiment and claimed that the proposed model could improve the recommendation accuracy.

In 2018, P Sushmitha Singh, et al., [28] briefs about the happenings in ATRS and PTRS and compared various methods adopted in them. They assessed all techniques such as Hybridization, ML, Fuzzy, and Heuristic Search used in these algorithms for explaining implementation, advantages, and disadvantages. As a concluding note, P Singh found that location based CF and related user mining techniques used in PTRS can improve the performance of ATRS.

Kori et al [29] discussed a method that focuses on the blogs to extract the travel routes of a particular user based on his past blog entries as it contains valuable information for potential travelers which may not be available on official websites. Popular travel routes and places can be correlated to present multimedia contents relevant to those routes. They introduced a sequential pattern mining technique for fetching the travel routes and PrefixSpan algorithm for fast mining of sequential patterns. The system can generate the most popular POI provided the data should be in a well-structured format. They used CaboCha method to analyze dependency and a Japanese Lexicon as a dictionary for action verbs. For conducting this experiment authors selected 16,142 blog entries and 74 major place names in Japan in May 2006.

Considering the time constraints of the traveller, Trip-Mine algorithm developed by E.H-C.Lu et al., could mine the optimal trip from a large number of attractions to user’s travel-time constraints based on their location [30]. The algorithm could address two important questions in planning a trip, such as evaluating the popularity of an attraction based on score value and estimation strategies of time required for the trip. For optimization of the algorithm in trip planning, they used three methods named Attraction Sorting, Low Bound-Checking, and Score Checking. With the help of extensive experimental evaluations, they proved Trip-Mine could deliver excellent results.
W.G.R.M.P.S. Rathnayake [31] introduced a Google Maps Based Travel Planning & Analyzing System (TPAS) to propose a solution for location-based travel recommendations to the user. TPAS can indicate a polarity of YES or NO to a travel plan by evaluating the route, distance, and weather conditions. Also, it could recommend an optimal route plan between source and destination including all interim destinations with climate details. The author used Quarter Circle Technique along with JQuery, AJAX and HTML. The online survey claims that there are about 64 million unique users globally uses Google Maps [32] which is less than one percent of the world population [33] and still, those users have not above feature in Google Maps.

Shelar S. et al., [34] in 2018 proposed a model to analyse the characteristics of existing travel packages and generate a thematic model of the Tourist Area Season Topic (TAST). By adding the latent relationships between tourists in each tour group, TAST model upgraded to TRAST. The relationship between travel groups and packages, travelers in different groups chooses the same package are also compared and evaluated. TouchMap, a cloud computing-based architecture developed by J Ching and team [35] could effectively recommend customized trips considering multiple constraints by user’s check-in activities. The method was developed to find the solution for reducing the computational cost by parallel cloud computing techniques [40]. The PTR system is designed with two functional components. Touch Map, a POI search module supported by HADOOP and Trip-Mine to efficiently plan the trip backed by Map Reduce techniques. The authors conducted rigorous experiments based on data crawled from Gowalla to evaluate the performance of Touch Map and concluded that PTR can produce outstanding results in terms of accuracy and efficiency.

In December 2019, Wafa and Yung [36] proposed a Travel recommendation mechanism with a motivation to pay more attention to Under-Emphasized Tourist Spots which are often unnoticed by tourists because of deficiency in the promotion. Apart from the conventional social media inputs like travel blogs, ratings, and reviews, and few latent factors like food, cleanliness, and opening hours are also considered for optimal recommendation. Latent Dirichlet allocation (LDA) algorithm for topic modeling and Support Vector Machine algorithm for the sentiment classification. It also used SentiWordNet lexicon-based sentiment analysis(SA) tool to collect sentiment of words. Rizwana.K.T et al., in 2019 [37] discussed the processing of NLP and SA with sophisticated algorithms to ensure precision and optimization. SentiWordNet designed to fetch input from all data sources and uses an artificial neural network (ANN) based learning model, the biggest challenge was processing the abundant amount of online data on different forums. The experiment is conducted by using 26 Korean travel websites, 94 locations, 1.4 lakh reviews, and more than 100 top relevant blogs. The authors claimed that the system could enhance the performance and achieve 94% prediction accuracy with state-of-art.

Valliyammai and Prasanna [38] proposed a model which can provide a high degree of the personalized recommendation of the most suitable travel destinations and the formation of like-minded travel groups based on users with similar interest. The conventional cold start and data sparsity problems of CF have been addressed in this system. Data classification is executed by Apriori algorithm and Fuzzy C Means algorithm. The commonly used keywords by travelers grouped to form a domain thesaurus which is useful in classifying the user review in given domains. The stemming method is adopted here to optimize and review processing. Text-based classifications are done by Naive Bayes classifiers. The three important segments in this model are Preference Keyword List Generation, Mining User Interest Patterns, and Travel Group Formation. While comparing different RS based on several criteria, Hybrid Recommendation methods can be further divided into tightly coupled and loosely coupled methods. CF is consid-

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Dataset Used</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cloud Computing and Hadoop</td>
<td>Gowalla</td>
<td>Low computation cost.</td>
<td>Dependent on cloud connectivity</td>
</tr>
<tr>
<td>2</td>
<td>Skyline Query</td>
<td>DEH System</td>
<td>Best POI recommendation along with the distance from the current location to POI.</td>
<td>Dependent on the connectivity of various interlinked servers</td>
</tr>
<tr>
<td>3</td>
<td>Hybridization</td>
<td>Gowalla</td>
<td>Efficient and accurate</td>
<td>Unstable Output</td>
</tr>
<tr>
<td>4</td>
<td>Heuristic Search</td>
<td>Ctrip</td>
<td>Extra information about mode of transport</td>
<td>Cannot recommend POIs in a new city effectively</td>
</tr>
<tr>
<td>5</td>
<td>Fuzzy Logic + Genetic Algorithm</td>
<td>Macau Map</td>
<td>A detailed itinerary with time duration for each POI visit</td>
<td>Not completely accurate recommendations</td>
</tr>
<tr>
<td>6</td>
<td>Hybridization &amp; Machine Learning</td>
<td>Trip advisor</td>
<td>Overcomes drawbacks of other methods and gives high accuracy</td>
<td>Gives only apt POIs, does not give a tour plan</td>
</tr>
</tbody>
</table>

Table 1. Performance Analysis of Travel Recommendation Methods
With the success of treating voice and text by using deep neural networks, AutoRec, for collaborative filtering, which works well for both explicit feedback and implicit feedback. To achieve the maximum benefit, NACE is designed to learn the hidden relationship between interactions that can reconstruct user/item preferences via a non-linear matrix factorization method. The design architecture of NCAE consists of four modules. An Input dropout module, a sparse forward module, a sparse backward module, and an error reweighting module. Qibing and team also designed a three-stage pre-training mechanism that combines supervised and unsupervised feature learning for optimizing the deep architecture of NACE. By conducting abundant experiments on three real-world datasets, the authors could prove NCAE consistently outperforms other methods by a large margin and significantly advance the state-of-the-art.

4. Discussion

The primary intention of this study was to analyze the development of recommendation systems in the travel domain, approaches and techniques being used, different data sources, the role of stakeholders, grey points, and challenges to be addressed. The analysis resulted in a classification of RS into three generations in connection with technological growth, stakeholders’ involvement, and quality of desired output delivered to the traveler.

As per the study, the first generation travel RS come mostly in the form of travel guides. The expected outcome of recommendation with relevant items to the user is attained by customization and personalization of these travel guides. The availability and classification of relevant data were one of the major hurdles at this stage. At the next stage, second-generation focused on reducing the overhead of travelers by reducing the explicit user inputs. The technology was matured enough to grab the required information implicitly from previous engagements of the user. Artificial intelligence and Mobile computing made these data analysis and prediction jobs easier and subsequently automatic identification of relevant data additionally turned out to be very easy. A broader view on RS came in picture in the third generation. It concentrates the recommendation in a multi-dimensional aspect by considering all solutions to address the tour plan of a traveler apart from destinations, points of interests, route plan, accommodation, food, budget, time, weather, mode of transport, etc. External sources, online activities, and social media generate Big Data, good enough to perform high performance computing and detailed analysis for more accurate recommendations to the users.

The summary of few of the discussed papers furnished in 2.
5. Conclusion

From the survey conducted in previous decades, the Travel recommendation system also matured along with the technologies. While the earlier travel recommenders could convey just static proposals with accessible inside information sources, the current frameworks could recommend exceptionally customized, real-time, context-aware suggestions. Internet-based life, Artificial Intelligence, and Collaborative algorithms could assume a vital job in producing personalized suggestions and meet the user requirements. A recommendation could be more popularised when it extends its functionalities to Local languages. The data analysis and prediction in Natural Languages will add a more personalized recommendation to any user. The investigation additionally recognized there is a requirement for more protection and security contemplations while structuring new recommender frameworks as the degree of personalization and contextualization are getting extended.

### References


Z. Zhang, Y. Morimoto, Collaborative hotel recommendation based on topic and sentiment of review comments, In: proceeding of the 9th forum for information and engineering, DEIM Forum, 2017.


D. Ajantha, Jobi Vijay, Raji Sridhar, A userlocation vector based approach for personalised tourism and travel rec-


**********
ISSN(P): 2319 – 3786
Malaya Journal of Matematik
ISSN(O): 2321 – 5666
**********