

A Survey on News Recommendation based on User Preferences and Location

Dhanashree More¹, Ajinkya Phand², Nagarjun Komarashetty³, Sampada Choudhari⁴, P.N.Vengurlekar⁵

^{1,2,3,4}BE Student, SITS, Pune

⁵Prof., Dept. of Computer Engineering, SITS, Pune, India

Abstract - Recommender systems are becoming an essential part of smart services. When building a news recommender system, we should consider special features different from other recommender systems. Hot news topics are changing every moment; thus, it is important to recommend right news at the right time. Reading the news is the favorite hobby for many people anywhere in the world. With the popularity of the Internet and social media, users are constantly provided or even bombarded, with the latest news around the world. With numerous sources of news, it has become a real challenge for users to follow the news that they are interested. In this work we follow two steps. First, to take the user's personal interest and their location preference respectively. We are taking the user's personal interest implicitly and explicitly. Second, by taking into consideration the above preferences, our system will map the news articles with respect to the location and provide them to the user. Further we are also providing dictionary as well as summarized news articles.

Key Words User Profile, User Location, News articles similarity function, News recommendation System

1. INTRODUCTION

NOWADAYS, news reading is an indispensable daily activity of many people. With their recent popularity of smart mobiles and the rapid development of the mobile Web, more and more people tend to read news online via their mobiles or other handheld devices, e.g., tablets. However, due to the huge volume of news articles generated every day, readers cannot afford to go through all the news online. So, news recommendation systems, which aim to filter out irrelevant online information and recommend to users their preferred news, have been widely studied [1], [2], [3], [4].

In classic personalized news recommendation systems, a user's news preferences are usually learned using his/her news reading history or other online activity histories; therefore, the user's news preferences are static and dynamic in these systems. For the dynamic user preferences, can be obtained in efficient way using user's Facebook profile respectively. So personalized news recommendations will recommend the news to the people based on their preferences. However, in real-world contexts, users' news preferences usually evolve with the change of their locations.

So, the solution is system to provide news based on location preferences of the user, so that user will get updated local news also. For the user's preferred location, Geolocation can be used to fetch the location with the help of its Latitude and Longitude. For further more additional features providing summarized news articles using Open NLP API's as well as providing a search dictionary to search a meaning of the word which user want to search can be a better option.

The rest of this paper is organized as follow. Section II presents related work. Section III describes our literature review of news recommendation based on user preferences and location. Section VI describes the Problem domain. Finally, Section V concludes.

2. RELATED WORK

Facebook also uses recommendation system that maintains a profile of user's and recommend appropriate user's interest, friend suggestions to the user respectively (For example: People you may know section). Another social media platform like YouTube/Netflix analyses user's previous watch history and recommend the relatable videos to the user.

Similarly, SysKill and Webert build a recommendation system that recommends web pages based on some importance, where the importance is calculated by weight which is determined by term frequency/inverse document frequency (TF-IDF) [1].

3. PROBLEM DOMAIN

The information overloading and irrelevant information extraction is major problem of today. Hectic schedule and poor knowledge of technology exhaust the user during information searching and information retrieval. This problem becomes more sensitive and crucial when we try to extract current affairs and news from newspapers and online sources

Everyone has difference perceptions and different reading liking. It may vary as per user preference and job requirement. Popularity of content and impact of information is also important for user search.

A News Paper divided into various sections like city, sports, editorial, international, national, entertainment etc. All this section has equal importance and different user followers. Some time there may be possibility that, they may consist relevant information but in different sections and different newspapers. News Recommendation System can overcome this problem and suggest relevant news according to user preference, popularity factor and also on user's location [6].

4. LITERATURE REVIEW

According to literature survey, after studying different IEEE paper, collecting some related papers and documents some of the point are discussed here:

Content Based News Recommendation System Based on Fuzzy Logic [1]. Published by Md. Nuruddin Monsur Adnan, Mohammed Rashid Chowdury, Iftifar Taz, Tauqir Ahmed, Rashedur M Rahman propose system in that they describe following:

Fuzzy logic is an approach that helps in computing based on "degrees of truth" rather than the Usual "true or false" (1 or 0) Boolean logic. There is a simple reason behind using fuzzy logic. Related or recommendable news articles are not easily translated into the absolute terms of 0 and 1. We cannot absolutely point out a certain article 'X' and say that it is related to 'Y'. That is why in this paper author tried to develop a fuzzy system from several attributes of a news article which will eventually describe whether an article is worth for recommendation to a user or not.

They collected news articles from a single website (bdnews24.com) and no user data was available with them so they concentrated purely on content-based recommendation. Single article was considered as single dataset for easier processing. The news content was gathered by a crawler created by the authors and then contents were saved into MySQL database. After the pre-processing, they identified all the important keywords like nouns, adjectives present in different parts of the article and appropriate priorities were assigned to the keywords. After this prioritization the team developed a function which would calculate the degree of similarity between two articles based on priorities of the words present in the article. Based on the similarity score for the articles the relevant news articles were suggested to users.

Recommending News Based on Hybrid User Profile, Popularity, Trends, and Location [2]. Published by Suraj Natarajan Melody Moh, describe the following:

In this paper, the main focus of the authors was on three things: hybrid user profile generation, temporal dynamics, ratio of popular vs. trendy news. The hybrid profile of the user was generated using three approaches, click through analysis: in this method the most visited URL of user was taken and categorized appropriately, User Tweet Analysis: in

this user's tweets were analysed to catch user's interest, User Twitter Follower: in this method user friend list was captured and categorized. Location Based recommendation was done by getting current (or another preferred location) of user. Current location was mapped with WOEID (where on earth ID) which would point out the exact location. Based on this user profile was generated for each user.

Later on, to get popular and trendy news author computed cosine similarity between popular tweets and news articles which measured frequency of common words appearing in both documents. User was also provided with facility to select ratio of popular vs trendy news. To evaluate the accuracy of the recommendation system, authors measured how much users spent on the news articles of their top 5 interests predicted by the system, and found the ratio, and therefore the accuracy, to be as high as 94%. The addition of user location preference, users were more satisfied with the news recommended, which was also been demonstrated by many more news clicks.

Location-Aware Personalized News Recommendation with Deep Semantic Analysis [3], Published by Cheng Chen, Xiangwu Meng, Zhenghua Xu, and Thomas Lukasiewicz describe the following:

In this paper the author mainly focuses on system called LP-ESA (Location-aware Personalized news recommendation with Explicit Semantic Analysis) that offers personalized news recommendation based not only on user's current location but also on user's personal interests.

The author in the paper consider user's browsing history for explicit analysis of user interests. Heuristic and model-based methods like cosine similarity, Euclidean distance and Bayesian network were used by the authors to compute the similarity between user preferences and news articles. For the location based recommendation GeoFeed[4] and GeoRank[5] were used by the author. GeoRank uses the static point location of both user's and news articles while GeoFeed uses spatial extent. Explicit Semantic Analysis (ESA) is a Wikipedia concept in which each location and news article is represented as Wikipedia-based topic vector.

ESA collects set of documents for each location based on geo-tags and forms a vector. Then cosine similarity is being used to map location, user profile and news categories. In order to further improve the performance of LP-ESA, author proposed LP-DSA i.e Location-aware Personalized news recommendation with deep semantic analysis that utilizes deep neural network and the rest mapping is done.

Personalized News Recommendation using Classified Keywords to Capture User Preferences [4], Published by Kyo-Joong Oh, Won-Jo Lee, Chae-Gyun Lim, Ho-Jin Choi describe the following:

In this paper the author mainly focuses on the model, based on deep neural network, to analyse user preference for news

recommender system. The model extracts interested keywords, to characterize the user preference from the set of news articles read by the user in the past. Firstly, the author 's collected the latest news set via implemented news crawler day to day. Also, they did same pre-processing for each article in the latest news set such as parsing, noun extraction, and TF-IDF is done. Finally, they could get a bag-of-words of articles and the importance weights of each word's similar user profile. After these pre-processing, the recommender calculated cosine similarity every bag-of-words of the articles with user profiles. Then, every article could get the similarity score. Based on the score, the recommendable articles are ordered by rank. The recommendation results were delivered to each user respectively.

5. CONCLUSIONS

Recommender systems have made significant progress in recent years and many techniques have been proposed to improve the recommendation quality. However, in most cases, new techniques are designed to improve the accuracy of recommendations, whereas the recommendation diversity has often been overlooked. Proposed system will not only observe the news content on user preference or popularity basis but also provides news according to the user's preferred location. Proposed system will also provide summarized news articles and dictionary for user's

Comfort and betterment. After performing the extensive literature survey and recognizing the various methods used by the various authors, we have proposed our system architecture and the basic design workflow of our system. For instance, we have selected Jsoup library for scrapping and HTML geolocation for fetching user's location. Proposed system will also help to refine popular and effective news content according to user desire.

Our future scope includes improving the performance of the system, generating analysis report of the news articles and keeping track of users who spend time on a particular category of news articles and learn accordingly.

ACKNOWLEDGEMENT

Authors want to acknowledge Principal, Head of department and guide of their project for all the support and help rendered. To express profound feeling of appreciation to their regarded guardians for giving the motivation required to the finishing of paper.

REFERENCES

[1] "Content Based News Recommendation System Based on Fuzzy Logic" (Md. Nuruddin Monsur Adnan, Mohammed Rashid Chowdury, Iftifar Taz, Tauqir Ahmed, Rashedur M Rahman).

[2] Recommending News Based on Hybrid User Profile, Popularity, Trends, and Location Suraj Natarajan Department of Computer Science San Jose State University San Jose, CA, USA surajnm@gmail.com, Melody Moh Department of Computer Science San Jose State University San Jose, CA, USA melody.moh@sjsu.edu 2016 International Conference on Collaboration Technologies and Systems.

[3] Location-Aware Personalized News Recommendation with Deep Semantic Analysis Cheng Chen, Xiangwu Meng, Zhenghua Xu, and Thomas Lukasiewicz 2016

[4] J. Bao, M. Mokbel, and C.-Y. Chow, "GeoFeed: A location-aware news feed system," in Proceedings of the IEEE International Conference on Data Engineering, 2012, 54–65.

[5] J. Bao and M. F. Mokbel, "GeoRank: An efficient location-aware news feed ranking system," in Proceedings of the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2013, pp. 184–193.

[6] Priyanka Jain, Savita Rathod, "A STUDY OF NEWS RECOMMENDATION SYSTEM: A REVIEW" International Journal of Research in Science & Engineering e-ISSN:2394-829 Volume: 3 Issue: 3 May-June 2017.