Research Paper
Volume : 5 | Issue : 8 | Special Issue August-2016 • ISSN No 2277 - 8179

Enhanced D-Mobifeed with Serendipity Feature

Engineering

KEYWORDS: Index Terms—LANF, D-Mobifeed, three stage heuristic algorithm, h-categories, diversity, serendipity.

S.Kavitha — Asso.Professor, Computer Science and Engineering Vedavyasa Institute of Technology Karad,Malappuram

Amritha PP — PG Scholar, Computer Science and Engineering Vedavyasa Institute of Technology Karad,Malappuram

ABSTRACT

This paper provides the latest technology involved in news feed system. A LANF system generates news feed for mobile users based on their spatial preference and non-spatial preference. Existing LANF systems do not consider the concept of diversity and only consider the relevancy and preference of news feed. D-Mobifeed is introduced to overcome the limitation of existing LANF system. This system enables users to post and generate geo-tagged messages as news feed and each news feed contain at least h different categories. Thus the total relevancy of news feed is maximized and diverse result is obtained for user. D-Mobifeed uses the three stage heuristic algorithm to achieve diversity, relevancy and efficiency of news feed. By using this algorithm over brute force algorithm, the efficiency improves for about an order of magnitude. The concept of serendipity is also included in the system so as to improve user satisfaction. Serendipity is happy or accidental surprise. User gets notification about his favorite spot while he passes through the location.

I. INTRODUCTION

Technology has developed to reach heights in a manner that people rely on machines to do their work in all aspects of life. Man has decided to choose machine over manual power from ages ago. Smart phones can be said as the greatest invention of science and technology which have made a drastic impact on modern era. The enhanced features that come along with smart phones are even jaw dropping at times. From simple calculation, calendar applications to mind reading and health check. The advent of smart phones made man to live in a time where smart phones have become inevitable to live in the modern society. Wireless communication and ubiquity of GPS have added benefits to the usage of smart phones. All of us are interested in news feeds that suggest the dimensions of a certain locality or an event. Location Based Social Networks (LSBN) have become more prevalent and location aware with the advent of wireless and GPS equipped smart phones. Facebook and 4share are some of the examples for LSBNs. One of the major role that these play is to send or receive a news feed. It enables a user to post or receive a geo-tagged message as news feed from anywhere, anytime.

Location Aware News Feed system consider user's current and future locations and intend to generate news feed based on their location. Many existing system like Geofeed and Mobifeed consider the most relevant geo tagged messages as news feed for users. Mobifeed have the intention to generate the news feed at user's current and future locations, whereas Geofeed intends to generate news feed that are relevant to current location and does not consider future locations of user. The disability to consider future location is a major drawback of Geofeed since the news feeds are not optimized. In order to improve the relevance of news feed at current and future locations Mobifeed have the GPS technique to calculate the future location. Though relevance is solved with Mobifeed there arises the concept of diversity which is meant to be the most essential for user satisfaction. Although users expect to receive messages that are highly relevant to their interests, they may prefer a location aware news feed with a certain level of diversity. That is the messages belong to certain number of categories. This work considers a mobile environment that makes the location- and diversity-aware news feed system unique and more challenging. With the geographical distance between a message and a mobile user in a relevance measure model, the relevance of a message to a mobile user is changing as the user is moving. Such a dynamic environment gives us an opportunity to employ location prediction technique to improve the quality of news feeds and the system efficiency.

Existing diversification problems focus on retrieving an individual list of items with a certain level of diversity. In contrast, with location prediction techniques, that aim at improving the quality of news feeds by scheduling multiple location- and diversity-aware news feeds for mobile users simultaneously. At this end the proposal of Dmobifeed plays a vital role. Dmobifeed is a framework that takes into account both relevance and diversity of news feed, where D stands for diversity. This framework is most beneficial for mobile users as they can schedule the news feed to their interest and with most relevancy. The concept of Dmobifeed was put into action with an interesting principle of h-diversity. H-diversity constrain was borrowed from the concept of l-diversity principle which is used to generalize non sensitive attributes like age, zip codes etc. These values are used to protect privacy of publishing data. It should be at least l different values similarly we have h different values for category. Thus enhancing the versatility of categories that user can rely on.

The major aspect that is followed in this framework are relevancy, diversity, optimization and finally serendipity. The relevancy of news feed is determined by the usage of Global Positioning System (GPS), where the exact location information is retrieved to mark the relevancy of news feed for user. Diveresity is achieved with the concept of h-diversity principle, where user get to experience the wide range of news feed from different categories. Optimizing the most relevant and desired news feed is again achieved. A new concept of serendipity is introduced to bring out a happy surprise for the user. Most choosen category is saved in users account and it is set as his preference. The preference once set and if user passes through the location where his favorite is set, he would get a notification about the newsfeed.

II. RELATED WORK

Extensive literature survey is done for a captive study about existing systems and the technology used to develop the proposed system. Highlighting the state of the art techniques in location aware news feed systems and existing diversity models in recommender systems and web search systems. This section presents some of the most recent research works related to this research area and possible solutions suggested by most eminent authors.

A. Mobifeed

Wenjian Xu, Chin-Yin Chow and Man Lung Yiu have proposed a Location aware news feed system known as Mobifeed [1]. A location aware news feed system enables a user to share geo-tagged user

generated messages, i.e. user can receive nearby messages that are most relevant to her. Mobifeed is a framework designed for scheduling news feed for mobile users. The major functionalities include location prediction, relevance measure and news feed scheduler. The location prediction function is used to predict a user's locations based on an existing path prediction algorithm. The relevance measure function is implemented by combining the vector space model with non-spatial and spatial factors to determine the relevance of a message to a user. The news feed scheduler works with the other two functions to generate news feeds for a mobile user at her current and predicted locations with the best overall quality. To ensure that MobiFeed can scale up to a larger number of messages, heuristic news feed scheduler is designed.

B. Geofeed

Jie Bao, Mohamed F. Mokbel, Chin-Yin Chow proposed the GeoFeed [2] system, which is a locationaware news feed system that provides a new platform for its users to get spatially related message updates from either their friends or favorite news sources. GeoFeed distinguishes itself from all existing news feed systems in that it enables users to post message with spatial extent rather than static point locations, and takes into account their locations when computing news feed for them. GeoFeed is equipped with three different approaches for delivering the news feed to its users, namely, spatial *pull, spatial push,* and *shared push.* The system takes design of a smart model for GeoFeed to decide about using these approaches in a way that: (a) minimizes the system overhead for delivering the location aware news feed, and (b) guarantees a certain response time for each user to obtain the requested location-aware news feed. GeoFeed also supports location-aware news feed function for its mobile users. Based on real and synthetic data obtained after various experimental results, show that GeoFeed outperforms existing news feed systems. Once a user u logs on to her favorite social network site that is equipped with GeoFeed, u will find the set of messages that are more relevant to her current location. This could be also said as the only system that provides spatial preferred news feed. The spatial filter allows this to happen.

C. Location Aware News Feed

Bao et al proposed location aware news feed [3] which deals on the principles of location aware news feed system. Most existing news feed systems only provide publish/subscribe services that simply forward messages to subscribed users. Injected the location-awareness into a news feed system, which enables a message to be associated with a spatial extent to control where users can receive it. The framework Mobifeed is proposed in such a way that it is designed to schedule news feed for mobile users. MobiFeed takes the limitations of mobile devices and the user's preferences into account, and schedules the most relevant geo-tagged messages to mobile users. Unfortunately, MobiFeed has a major limitation that only considers the relevance of messages to users, so a news feed may contain messages related to the same category; and thus it would impede users to discover new places and activities. In conventional web search/recommender systems, topic diversification is a key method to improve user satisfaction. In order to eradicate this limitation it is brought in idea to implement the concept of diversity and thus paved the way for D-Mobifeed.

D. L- Diversity Principle

A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam proposed l-diversity: Privacy beyond k- anonymity [4]. The l-diversity principle is proposed for privacy preserving data publishing. Basically, this principle is used to generalize non-sensitive attributes (e.g., zip codes 13053 and 13068 are generalized to "130**" and ages 28, 29, and 21 are generalized to "< 30") in a class of records such that the sensitive attribute contains at least l different values, in order to protect the privacy of published data. The entropy l-diversity is further used to defend against the homogeneity problem without considering the role of background knowledge, i.e., entropy increases as frequencies of sensitive attributes become more uniform. In this work, it focus on a different problem because D-MobiFeed aims to maximize the relevance of news feeds for mobile users while news feeds satisfy the h-diversity constraint (i.e., the messages in each news feed belong to at least h categories). Considering D-Mobifeed the h diversity concept is borrowed from the l-diversity principle. H is just a random number to denote the number of categories.

E. Toward The Next Generation Of Recommender Systems

Recommender systems are gaining importance in today's life. G. Adomavicius and A. Tuzhilin proposed towards the next generation recommender systems [5], where the overview of the field of recommender systems and describes the current generation of recommendation methods that are usually classified into the following three main categories: content-based, collaborative, and hybrid recommendation approaches. The limitations of current recommender systems and methods to overcome it are put together in this paper. Recommender systems have become an important research area. Many applications use recommender systems to improve their service and some of them are amazon, movielens etc

F. Improving Recommendation Lists Through Topic diversification

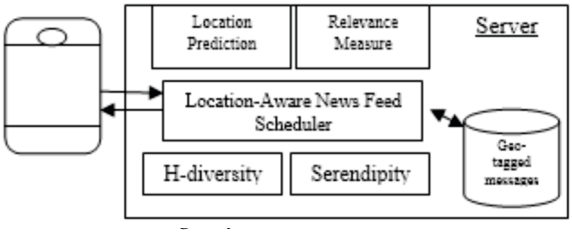
Topic diversification [6], an algorithmic framework to increase the diversity of a top-N list of recommended products. In order to show its efficiency in diversifying, also introduced our new intra-list similarity metric. Contrasting precision and recall metrics, computed both for user-based and item-based CF and featuring different levels of diversification, with results obtained from a large scale user survey, showed that the user's overall liking of recommendation lists goes beyond accuracy and involves other factors, e.g., the users' perceived list diversity.

G. Using Twitter to Recommend Real-Time Topical News

RSS (Really Simple Syndication) and Twitter are two important Web 2.0 technologies. The former is a data format that is designed to provide access to frequently updated content. Most commonly, RSS is used as a way to syndicate or distribute news information in the form of short-updates that can be linked back to complete stories. RSS Readers then allow users to aggregate the updates from many different feeds to provide a one-stop-shop to breaking news, although as users subscribe to tens of RSS feeds this introduces a niche information overload problem. Twitter in contrast is a so-called micro-blogging service. It allow users to submit their own short status update messages, called tweets, while following the status updates of others. Recently there has been much interest in Twitter, partly because of its growth and partly because of its ability to provide access to thoughts, intentions and activities of millions of users in real-time. A prototype system has been developed and deployed and early evaluation results suggest that users do benefit from the recommendations that are derived from the Twitter data.

III. SYSTEM ARCHITECTURE

The figure 1 below depicts the system architecture of DMobifeed. The framework of Dmobifeed is adapted from Mobifeed with additional concepts of h-diversity and serendipity added. The major entities are Category-associated geo-tagged messages & system users.



Category-associated geo-tagged Messages: The message collection is denoted with M, where each message belongs to a tuple (Message_ID, Sender _ID, Content, Timestamp, Spatial, and Category). Message ID is unique. Each news feed is unique and to denote it, Message_ID is used.

Sender_ID is unique for a sender. He uses a unique ID to post his news feed. Content is the whole news feed, which the user will type down. Content gives the review about a certain place or event. Timestamp includes the time and date of the post which is used in calculation of relevancy. Spatial involves the latitude and longitude of a locality and also includes the address of a locality. Final entity in the tuple is category, which is the most important. Category based classification of news feeds are done in the system. News feed from minimum of h different category is obtained for a user.

System user: User can make use of the system by using a GPS equipped mobile phones. The user is the source of news feed. He can post the news feed, which is his opinion about a certain location, a specific categorized place or an event. These news feed may be useful for others who may be looking for reviews about a certain location or event. While posting the news feed the user's location information & date and time is retrieved from his mobile phone. With the Global Positioning System technique the latitude, longitude and address of a locality is known. The message category and the news feed are added along by the user. The messages that are posted by the user are set in the system as news feed. The user can send a query to retrieve the news feed. These news feed will be relevant to his locality. The relevancy is measured by using the latitude and longitude of a locality. The latitude and longitude captured during the posting of the news feed and latitude and longitude while querying is calculated and the ones with least difference are the most relevant news feed.

A. Location Prediction

Any location prediction algorithm can be employed to find the location of user. The location prediction function helps in finding the location at the present and future. The latitude and longitude differences of current and past location are calculated. The difference is then added to current location and the range of distance is noted. By this way the future location range of a user can be calculated.

B. Relevance Measure

The relevancy of news feed is important in location aware news feed system. The news feed that are obtained as a result of querying should be useful and most relevant to the user. The technique of GPS plays a vital role in helping this out. User's current location is calculated and news feed matching this location is only provided to the user. Timestamp also is utilized at this stage to improve the relevancy of news feed.

C. Serendipity

Serendipity is an accidental surprise or happy surprise. The concept of serendipity is added to the system to improve the user satisfaction. The users that rely on the system are moving users. Querying and driving is not possible. User's favorite category is calculated based on the number of counts he have selected it. The count is incremented each time the user selects the item. While user passes through his favorite spot, he receives a notification message in his mobile phone. He can either discard the message or click on it and the message will read out.

D. H-diversity

The most important concept is diversity. Users are not only interested in relevancy of news feed they are always looking for diversity. H-diversity concept is borrowed from l-diversity principle. Here the value of h is the total number of categories available in the system. If the user is not being choosy, he can get a list of all items with diversified result. Suppose there are 10 categories or the value of h equals 10, then the user will receive news feed among the 10 categories, provided the data is present in that location.

IV. H-DIVERSITY CONSTRAIN

The h-Diversity Constraint Checking problem is nontrivial because a brute-force method has to try all possible combinations of news feeds. The scheduling step receives the minimum total diversity from the decision step. In this section, we redefine the requirement of the h-Diversity Constrained Scheduling Problem. The main objective of the problem is to calculate the minimum diversity among all the news feed with relevancy, uploaded by the user. To solve the problem brute force algorithm was employed, but the method is very costly for solving the problem. To this end a three-stage heuristic algorithm was proposed.

A. Three Stage Heuristic Algorithm

The algorithm is implemented in three stages to resolve the h diversity constrain. In the first stage the relevancy of the news feed is calculated. This is a process of filtration, where the relevant news feed for the user is obtained. Among the relevant news feed the second stage process is implemented. In the second stage h diversity constrain is checked. The data obtained in the first stage is passed through the h diversity constrain. The data obtained in second stage will contain news feed that belong to at least h different message category. Final stage of the algorithm is scheduling the news feed. The final result is provided to the user once he makes a query. The most relevant and diverse result is obtained with three stage heuristic algorithm

V. CONCLUSION

DMobiFeed; a location-aware news feed framework takes the relevance and diversity of news feeds into account when scheduling news feeds for moving users DMobiFeed users can specify the minimum number of categories in a news feed as an h-diversity constraint, and it aims at maximizing the total relevance of generated news feeds and satisfying the h-diversity constraint. The two key problems in DMobifeed is decision problem and optimization problem. The decision problem is solved using the strategy that whether h diversity constrains can be fulfilled or not. For the optimization problem an effective three stage heuristic algorithm is used. Thus the relevant and diverse result is obtained for the user who query. The concept of serendipity is an enhanced feature for DMobifeed. This feature is effective for mobile users and it improves user satisfaction.

REFERENCES

[1.] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[2.] J. Bao, M. F. Mokbel, and C.-Y. Chow. GeoFeed: A location-aware news feed system. In IEEE ICDE, 2012

[3.] R. Agrawal, S. Gollapudi, A. Halverson, and S. Leong. Location aware system.. In ACM WSDM, 2009.

[4.] Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam.l-diversity:. ACM TKDD, 1(1):3, 2007.

[5.] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions.

[6.] M. Zhang and N. Hurley. Avoiding monotony: Improving the diversity of recommendation lists. In ACM RecSys, 2008.

[7.] O. Phelan, K. McCarthy, and B. Smyth. Using twitter to recommend real-time topical news. In ACM RecSys, 2009.