IIMB Management Review (2016) **28**, 25–30

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/iimb

IMR-DCAL SPECIAL PAPER

A study and analysis of recommendation systems for location-based social network (LBSN) with big data

Murale Narayanan [a,*], **Aswani Kumar Cherukuri** [b]

[a] *Information Technology, EMC Corporation, India Center of Excellence, Bangalore, India*
[b] *School of Information Technology & Engineering, VIT University, Vellore, Tamil Nadu, India*

Received 28 March 2014; revised 15 June 2015; accepted 22 January 2016; available online 13 February 2016

KEYWORDS
Big data;
Data mining;
Recommendation system;
Social network;
LBSN

Abstract Recommender systems play an important role in our day-to-day life. A recommender system automatically suggests an item to a user that he/she might be interested in. Small-scale datasets are used to provide recommendations based on location, but in real time, the volume of data is large. We have selected Foursquare dataset to study the need for big data in recommendation systems for location-based social network (LBSN). A few quality parameters like parallel processing and multimodal interface have been selected to study the need for big data in recommender systems. This paper provides a study and analysis of quality parameters of recommendation systems for LBSN with big data.

Introduction

Recommender systems or recommendation systems (RS) collect information based on the preferences of users (for example—songs, movies, jokes, books, travel destination and e-learning material). Recommender systems work based on users' information from different sources and provide recommendation of items. This information can be explicit (user rating) and implicit (monitoring user's behaviour), with millions of users using social networking services like Facebook, Twitter, and so forth. The rich knowledge that has accumulated in these social networking sites enables a variety of recommendation systems for its users.

A social network is an abstract structure comprised of individuals connected by one or more types of relations, such as friendship, shared knowledge, and common interests as stated by Zheng, Zhang, Xie, and Ma (2009). Location data add strength to the connection of the social networks. A location can be represented in relative, absolute, and symbolic form. Location is usually represented in three kinds of geographical representations—a point location, a region, and a trajectory.

In recent times, localisation techniques have enhanced social networking services, allowing the users to share their location-related content, and locations such as geo-tagged

* Corresponding author. Tel. +91 9972565848; fax: 41659511.
E-mail address: muralen@gmail.com (M. Narayanan).
Peer-review under responsibility of Indian Institute of Management Bangalore.

http://dx.doi.org/10.1016/j.iimb.2016.01.001

photos and notes. This is known as location-based social networks (LBSNs) (Zheng et al., 2009). An LBSN adds a location to an existing social network, and also tells the people in their social network that they can share their location-related information. Based on the location-related information, a new abstract structure is derived and connects connected individuals based on their location-related content, such as photos, texts and videos. Instant location and the history of a person are given as a timestamp during a certain period.

The advances in wireless communication technologies and location acquisition enables people to add a location dimension to traditional social networks and promotes a bunch of LBSN services, such as Foursquare, GeoLife and Loopt, where users can easily share their experiences in the physical world through mobile devices. The location dimension bridges the gap between the physical world and the digital online social networking services, giving rise to new opportunities and challenges in traditional recommender systems in the following aspects—complex objects and relations, and rich knowledge.

Location is one of the important components of user context and implies extensive knowledge about a user's interests and behaviour, thereby providing us with opportunities to better understand users in an abstract structure not only according to user behaviour, but the mobility of the user and his/her activities in the physical world. In recent times location-based services, such as tour guide and location-based social network, have accumulated a lot of location data. Today, the positioning function in mobile devices, such as GPS-phones, lets people know their locations easily. This location data provide various location-based services on the web and has shown itself to be attractive to the users. In real time, data are huge in volume, but data warehouses use small-scale datasets of users for recommendation.

When it comes to real-time scenario, these techniques may fail because millions of users will use social networks at the same time. The major challenges to be addressed in LBSN recommendation are 1) location-context awareness; 2) heterogeneous domain and 3) rate of growth.

Different types of data sources are used in recommendation systems for LBSNs, including 1) user profiles, 2) user online histories and 3) user location histories. This involves huge volumes of data in real-time scenario. Most recommendation systems in LBSNs currently use only one type of data source to make recommendations. Moreover, many of the data sources are related and may mutually reinforce each other. By considering more diversified data sources, more effective recommendations can be provided. For instance, the user online interactions, social structures and location histories are all very relevant to friend recommendation. If two users have more online interactions, are close in the social structure, and have overlapped location histories, these users are likely to be compatible. A friend recommender system that can consider all these factors will make higher quality friend recommendations.

We carried out an analysis, based on the characteristics of a recommender system, to give a comparison between big data and data warehouse with a dataset collected from Foursquare users, using a qualitative approach. This paper is organised as follows: Section 2 deals with literature review; Section 3 explains the challenges of the domain; Section 4 provides the objective of the paper; Section 5 details the dataset discussed in this paper; Section 6 gives characteristics of a location-based recommendation system; Section 7 explains the qualities of the location-based recommendation system, and Section 8 provides the conclusion.

Related work

Social media recommendations

Social media recommendation aims to provide users with suggestions of photos, videos, or other web content they might like. Using location information in LBSNs can improve both the effectiveness and efficiency of traditional social media recommendations. Several works in spatial keyword search for web content show the effectiveness of this pairing (Bouidghaghen, Tamine, & Boughanem, 2011; Cao, Cong, & Jensen, 2010b, 2011; Chen, Geyer, Dugan, Muller, & Guy, 2009; Zhang, Chee, Mondal, Tung, & Kitsuregawa, 2009). Location-aware image ranking algorithms have been proposed to increase the relevance of search results (Arase, Xie, Duan, Hara, & Nishio, 2009; Kawakubo & Yanai, 2011; Silva & Martins, 2011), which in turn improves the quality of the image tags, using a recommender system to automatically infer and suggest candidate location tags (Daly & Geyer, 2011).

The efficiency of recommendation systems can be significantly improved by using location data to prune out irrelevant information (Scellato, Mascolo, Musolesi, & CrowCroft, 2011). This improves the efficiency of content delivery networks using a novel caching mechanism based on geographic location. A real-time recommendation system, as suggested in Sandholm and Dung (2011), has been built for online web content using a collaborative filtering method to make more diverse and personalised recommendations within a geographical area. Levandoski, Sarwat, Eldawy, and Mokbel (2012) have proposed a novel location-aware recommendation system (LARS) framework to exploit users' ratings of locations using a technique that uses the distance of querying users to influence recommendations.

Categorisation by methodology

Although traditional recommendation systems have been successful by using community opinions, such as inventories in Amazon (Linden, Smith, & York, 2003) and news from Google (Das, Datar, Garg, & Rajaram, 2007) incorporating location information requires novel approaches. In this section, we categorise the major methodologies used by recommendation systems in location-based social networks as being based on — 1) content-based recommendation and 2) link analysis.

Content-based recommendations

Content-based recommendation systems, such as context aware and location based using Bayesian model (Park, Hong, & Cho, 2007; Ramaswamy et al., 2009), match user preferences discovered from users' profiles with features extracted from locations, such as tags and categories, to make recommendations. These systems require accurate and structured information for both the user profiles and the location features to make high-quality recommendations. The major advantage of the content-based approach in such a system is that it is robust against the cold start problem for