

A Survey on Micro Video Recommendation System Based on Big Data

Suman Kotin¹, Prasadgouda Patil²

¹*Master of Technology (Computer Science), Secab Institute of Engineering and Technology, Vijayapur, Karnataka, India*

²*Asst Professor, Secab Institute of Engineering and Technology, Vijayapur, Karnataka, India*

Abstract- With the refinement in the Internet and long range social correspondence advantage, the video is ending up more recognizable, especially for younger's. In any case, for a couple of customers, they contribute a massive measure of centrality to get their most regarded records from aggregates accounts on the Internet; for the video producers, they don't fathom what sorts of watchers like their things. Thusly, this paper proposes a video recommendation system. The suggestion figuring's the explanation behind meeting of this structure. Standard recommendation figuring's join substance based recommendation, empowered exertion proposal estimations, and so on. At the Big Data times, the weights what we meet are data scale, execution of figuring, and specific perspectives. Along these lines, this paper restores the standard suggestion estimations, using the outstanding Neural Association Algorithm to process the Big Data.

I. INTRODUCTION

Video is a new form of information media. With the development of the Internet, 3G (the 3rd Generation mobile communication technology), and 4G (the 4th Generation mobile communication technology) network, the bandwidth and speed of network become faster and faster. These technologies provide conditions for dissemination of information media. Video is a short time video [1] [2], which lasts for 30 seconds to 300 seconds. The short time micro-videos are popular with young people, because the teenagers prefer to watch the micro-video on their comfortable time through mobile devices. For micro-video producers, the problem is they do not know how many people like their products, and do not know how many times their video have been watched. Therefore, this paper proposes a video recommendation system (MRS). One of the purposes is an overview of videos for the producer. In this

way, the producer knows how many users love their video, and how many times their videos are on-demand. Another purpose is for users. The system can analyze the users' favorites and watching history, automatically push appropriate video to the users. It is becoming more popular with the internet technology development, which means the data sets whose size is beyond the ability of current technology, method and theory to capture, manage, and process the data within a tolerable elapsed time [3]. In order to enhance the MRS accuracy, we need to collect large volume data sets about who and when watched the micro-video, how many times the micro-video on demanded, and how many people love the micro-video. Therefore, the MRS, proposing in this paper, use technology to process the collected data sets. Data sets are the foundation of the recommendation system. The first step of video recommendation is to collect data as far as possible from the Internet. We download data from video websites, video forum, video online chat websites, and so on. Web crawlers [4] [5], one of basic data sets collection, can download resources from Internet. The web crawlers originally used for search engine. In this paper, the results of crawler directly affect the accuracy of recommendation system.

II. LITERATURE SURVEY

1. Recommendation of YouTube Videos

YouTube is a huge video-sharing service with hundreds of millions of users and hundreds of thousands of videos being uploaded every day. Thus, recommendation of YouTube videos to a single user is a challenging problem which cannot be solved by simply reusing the prevailing recommendation methods. The paper presents a specific

recommendation algorithm for YouTube which relies on the data retrieved through the YouTube Data API. A cloud-based application integrates the proposed algorithm and offers a web interface to end users. The paper presents a preliminary analysis of the recommendation quality and lists YouTube Data API limitations which influence the design of recommender systems for YouTube videos.

2. The Video Recommendation System Based on DBN

Video recommendation system provides users with suitable video for users to choose, which is an effective way to get a higher user satisfaction and user stickiness. Therefore, video websites pay much attention to it, as well as scholars. The existing recommendation algorithms are fused machine learning algorithms to video recommendation system. Such as some studies the SVM algorithm combined with a recommendation algorithm based on content, or uses the BP neural network combined with collaborative filtering algorithm, to improve the algorithm accuracy. With the rapid development of Machine Learning, progresses in Deep Learning are considerable. Especially after the RBM training efficiency matter has been solved by the random sample, the reliability of multi-layer neural network is more clearly, also caused academic interest in depth of the neural network research.

3. Cloud-Assisted Video Recommendation System

Various personalized recommendation systems have been proposed. However, they ignore that the accelerated proliferation of social media data has led to the big data era, which has greatly impeded the process of video recommendation. In addition, none of them has considered both the privacy of users' contexts (e.g., social status, ages and hobbies) and video service vendors' repositories, which are extremely sensitive and of significant commercial value. This paper proposes a cloud assisted differentially private video recommendation system based on distributed online learning. In our framework, service vendors are modeled as distributed cooperative learners, recommending videos according to user's context, while simultaneously adapting the video-selection strategy based on user-click feedback to maximize total user clicks (reward).

4. A Micro-video recommendation system using improved slope one algorithm based on Big Data

This paper has proposed a micro-video recommendation system. The slope one scheme algorithm is based on rating of the users, and it is simple, efficient, easy to implement. However, the slope one scheme suffers from both new item problems and data sparsity, which affect the performance of recommender systems. To overcome the drawbacks of slope one scheme, this paper proposed an improved slope one algorithm which contain collaborative filtering algorithm. The slope one scheme is improved by introducing content similarity computation to overcome the new item problem. In item-based collaborative filtering algorithms, the target users rating to the target item can be predicted based on the ratings that the target user has rated and the content similarities of items. And clustering algorithm is used to overcome the problem of data sparsity. By combining the set of items into several different clusters based on the item rating data, the target users rating to the target users rating to the target item can be predicted based on which cluster the target item belongs to.

5. Focused Web Crawling Algorithm

Information greatly influenced our lives, our lifestyle and way of thinking. A web search engine is a complex multi-level system that helps us to search the information that available on the Internet. A web crawler is one of the most important parts of the search engine. It's a robot that systematically browses and indexes the World Wide Web. A focused web crawler is used crawling only web pages that are relevant to the user given topic or web page link. A focused

We will review most effective focused web crawling algorithms that determine the fate of the search system.

6. Mobile Recommender Systems

Mobile phones are becoming a primary platform for information access and when coupled with recommender systems technologies they can become key tools for mobile users both for leisure and business applications. Recommendation techniques can increase the usability of mobile systems providing personalized and more focused content, hence limiting the negative effects of information

overload. In this paper we review the major issues and opportunities that the mobile scenario opens to the application of recommender systems especially in the area of travel and tourism. We overview major techniques that have been proposed in the last years and we illustrate the supported functions. We also illustrate specific computational models that have been proposed for mobile recommender systems and we close the paper by presenting some possible future developments and extension in this area.

7. Recommender System Application Developments: A Survey

This paper reviews up-to-date application developments of recommender systems, clusters their applications into eight main categories: e-government, e-business, e-commerce/e-shopping, e-library, e-learning, e-tourism, e-resource services and e-group activities, and summarizes the related recommendation techniques used in each category. It systematically examines the reported recommender systems through four dimensions: recommendation methods (such as CF), recommender systems software (such as BizSeeker), real-world application domains (such as e-business) and application platforms (such as mobile-based platforms). Some significant new topics are identified and listed as new directions. By providing a state-of-the-art knowledge, this survey will directly support researchers and practical professionals in their understanding of developments in recommender system applications.

8. Survey on recommendation system methods

In recent years recommendation systems have changed the way of communication between both websites and users. Recommendation system sorts through massive amounts of data to identify interest of users and makes the information search easier. For that purpose many methods have been used. Collaborative Filtering (CF) is a method of making automatic predictions about the interests of customers by collecting information from number of other customers, for that purpose many collaborative base algorithms are used. CHARM algorithm is one of the frequent patterns finding algorithm which is capable to handle huge dataset, unlike all previous association mining algorithms which do not support huge dataset. This paper covers different techniques which are used in recommendation system and also proposes a new

system for efficient web page recommendation based on hybrid collaborative filtering i.e. using collaborative technique and CHARM algorithm which are coupled with the pattern discovery algorithms such as clustering and association rule mining.

9. Big Data Analysis: Recommendation System with Hadoop Framework

Recommendation system provides the facility to understand a person's taste and find new, desirable content for them automatically based on the pattern between their likes and rating of different items. In this paper, we have proposed a recommendation system for the large amount of data available on the web in the form of ratings, reviews, opinions, complaints, remarks, feedback, and comments about any item (product, event, individual and services) using Hadoop Framework. We have implemented Mahout Interfaces for analyzing the data provided by review and rating site for movies.

III. PROBLEM FORMULATION AND PROPOSED WORK

Existing System

Many users, they spend a lot of time to get their favorite videos from amounts videos on the Internet; for the video producers, they do not know what kinds of viewers like their products. Traditional recommendation algorithms include content-based recommendation, collaboration recommendation algorithms, and so on. At the Big Data times, the challenges what we meet are data scale, performance of computing, and other aspects. In the existing system YouTube is a huge video-sharing service with hundreds of millions of users and hundreds of videos being uploaded every day. Thus, recommendation of YouTube videos to a single user is a challenging problem. There is no cluster formation between the multiple users.

Problem Statement

The recommendation of YouTube videos to a single user is a challenging problem. There is no cluster formation between the multiple users. To overcome from single user recommendation by taking the

consideration of Neural association cluster we can make the cluster based on category.

Proposed System

The proposed system to give recommendation for the user, this means that if the user sees some videos in YouTube, we are taking the dataset of complete users and doing analysis based on neural Association cluster. So this makes cluster among the users, cluster based on the category, if we selected Singers attribute it checks in the cluster how many are associated with that singer. If they match any singers with that it gives recommendation to others.

Advantages of Proposed System

- Recommendation of associated singers names very easy.
- Easy to find the common information between multiple users using clustering method.
- Possible Outcomes:
- The possible outcomes are association with the particular singer in the cluster based on the category.
- Limitations:
- Based on the existing dataset the recommendation is given to user.

Data Flow Diagram:

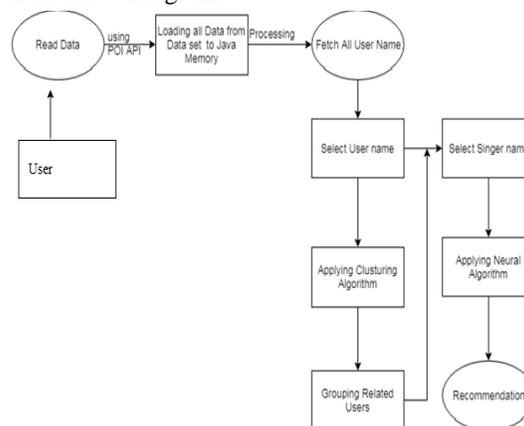


Fig: Data Flow Diagram

Sequence Diagram

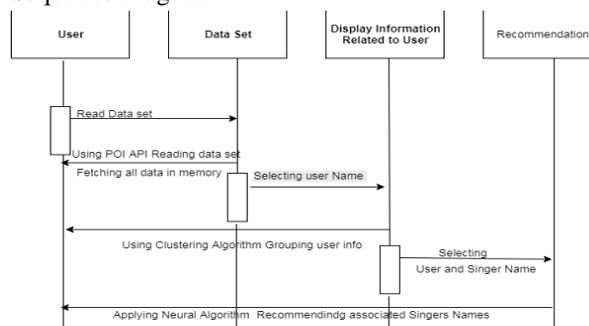


Fig: Sequence Diagram

IV. THE DESIGN OF SYSTEM STRUCTURE

With the development of the new generation of information network technology represented by mobile Internet, big data, cloud storage, the traditional advertising micro video creation has a fundamental change. This project studies the design of the IT architecture and the division of the system service level in the era of big data, and designs the structure of the system, which is shown in Fig.1:

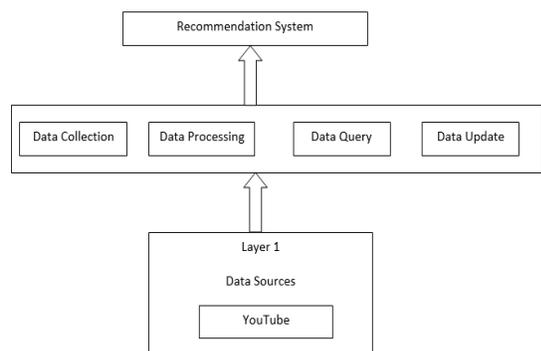
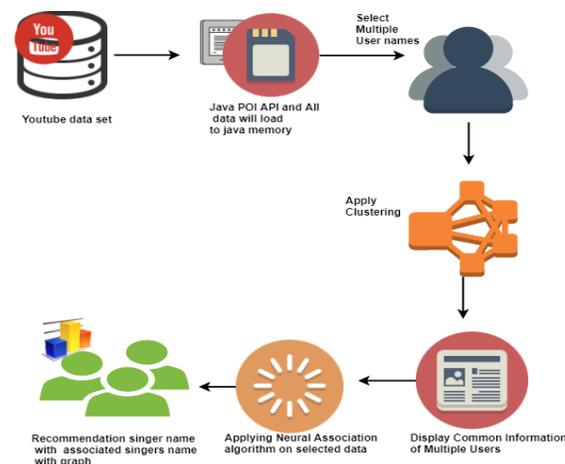


Fig: The structure of the system

V.SYSTEM ARCHITECTURE



Modules:

1. Reading dataset:

In the reading dataset first we need to read the dataset of youtube video links using Java POI API after

successful reading of dataset we need to load all the data into Java memory.

2. Selecting multiple users:

After successful reading of dataset we need to select the multiple users. After successful selection of the multiple users the corresponding video links will be fetched from the java memory.

3. Clustering:

After fetching the video links of the corresponding user the following steps we need to follow to form the cluster

Neural Association Clustering
NAC is a framework in this research which is interdependent among 3 blocks of flow.

Blocks:

- Neural
- Association
- Clustering

VI. CONCLUSION

With the change, the video are powerfully customary, particularly vivacious adolescents are doubtlessly going to watch accounts on cell phones. The inconveniences what we require at show is the strategies by which to locate the most venerated video. Something special, for video makers, what they control to is the thing that number of watchers like what sort of annals. In context of this see, this paper proposes a video suggestion structure. As per the watchers' looking at or watching history, this structure can support the most revered annals to the watchers. Obviously, this structure can collect the responses and give a few suggestion for video makers with what number of watchers like the video. The center furthest reaches of the framework is the recommendation figurings. The all things considered Neural Association Cluster estimations are suite for custom date sets, for example, content-based suggestion, encouraged effort recommendation, et cetera. Notwithstanding, with the movement of Big Data, the suggestion estimations ought to have the ability to manage the Big Data. Plus, our framework will give suggestion for related vocalists names and records.

REFERENCES

- [1] Y. Z. Li, T. Gao, and X. Y. Li, "Design of video recommender system based on cloud computing", *Journal on Communications*, Vol. 34, No. Z2, pp. 138-140, 147, 2013.
- [2] Y. Li, "Development mode of micro-video communication", *Academic Exchange*, Vol. 248, pp.177-181, 2014.
- [3] S. M. Meng, W. C. Dou, X. Y. Zhang, et al., "KASR: a keyword-aware sevice recommendation method on MapReduce for Big Data application", *IEEE Transactions on parallel and distibuted system*, Vol. 25, No. 12, 2014.
- [4] D. M. Zhou, Z. J. Li, "Survey of high-performance web crawler", *Computer Science*, Vol. 36, No. 8, pp.26-29, 53, 2009.
- [5] G. Y. Su, J. H. Li, and Y. H. Ma, et al., "New focused crawling algorithm", *Journal of Systems Engineering and Electronics*, Vol. 16, No. 1, pp.199-203, 2005.
- [6] X. W. Meng, X. Hu, and L. C. Wang, et al., "Mobile recommender system and their applications", *Journal of Software*, Vol. 24, No.1, pp. 91-108, 2013
- [7] G. X. Wang, H. P. Liu, "Surevey of personal recommendation system", *Computer Engineering and Applications*, Vol. 48, No. 7, pp. 66-76, 2012.
- [8] G. F. Sun, L. Wu, and Q. Liu, et al., "Recommendations based on collaborative filtering by exploiting sequential behaviors", *Journal of Software*, Vol. 24 No. 11, pp.2721-2733, 2013.
- [9] L. Guo, J. Ma, and Z. M. Chen, et al., "Incorporating item relations for social recommendation", *Chinese Journal of Computers*, Vol. 37, No. 1, pp. 218-228, 2014.
- [10] Y. R. Wang, M. Chen, and H. H. Wang, et al., "A content-based filtering algorithm for scientific literature recommendation", *Computer Technology and Development*, Vol .21, No. 2, pp. 66-69, 2011.
- [11] W. F. Liu, J. Z. Gu, and X. Lin, et al., " A Big Data management and analysis system based on Hadoop and Mahout", *Computer Application and Software*, Vol. 32, No. 1, pp. 47-50, 2015.
- [12] E. Jain, S. K. Jain, "Categoring Twitter users on the basis of their interests using Hadoop/Mahout

- platform”, Industrial and Information Systems (ICIIS), 2014 9th International Conference, pp. 15-17, 2014.
- [13] B. Fan, C. Jia, “Visual framework for big data in d3.js”, Electronics, Computer and Applications, pp.47-50, 2014.
- [14] Y. Z. Wang, S. Mao, “A blocks placement strategy in HDFS”, Computer Technology and Development, Vol. 23, No. 5, pp. 90-92,96, 2013.
- [15] X. S. Zhang, H. L. Wang, “AJAX Crawling Scheme Based on Document Object Model”, Computational and Information Sciences (ICCIS), pp. 1198-1201, 2012.