Volume 5, No. 4, April 2014 (Special Issue)
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

A New Strategy For Video Recommendation

B. Sudheer Kumar[1], A.Swetha[2], Ch.Deepa[3], Ch.Naveen[4], Y.Divya[5]
IT Department
ACE Engineering College, Hyderabad, India.

Abstract: Hadoop is a open-source software framework for storage and large scale processing. Hadoop Distributed File System (HDFS) is a primary storage system for large database. By using HDFS we ar e emphasizing on the video recommendation engine, the world's most popular online video community. The system recommends personalized set of videos based on their user requirements. Along with this, we recommend the frequently visited videos by the users as their suggestions by considering recent history. In addition to this, we will also show the user requested video(which was once not available when it was requested) when it gets available.

Keywords— HDFS, Recommendation,Absorption.

I. INTRODUCTION

A. Context and motivation of the project:

The rapid growth of the number of videos on video recommendation engine made it difficult to find videos of one's personal interest. Unfortunately, the difficulty of searching videos, made the size of the repository of new content as a daunting task. We can create an effective video suggestion system that depends on the analysis of the underlying videos. The whole idea is to create a personalized page of video recommendations that shows the latest most popular videos, recommendations tailored to their viewing habits. . In addition to this, we will also show the user the requested video(which was once not available when it was requested) when it's available.

II. EXISTING SYSTEM

Hadoop provides the path to read/write binary files. Anything that can be converted in to bytes can be stored in HDFS(videos). To add a file in to HDFS Hadoop uses Sequence Files. Sequence File is a flat file that consisting of binary key/value pairs. Sequence Files provides a Reader, Writer and Sorter classes for reading, writing and sorting respectively. By this we can convert video file in to sequence files and store in to the HDFS. After adding files in to HDFS to retrieve a file Map Reduce Algorithm is used. In Map Reduce we have both Key, Value for input as well as output. Here the video file gets split into many chunks, each Mapper process needs to interpret the bit stream chunk to provide access to the individual decode video frames for analysis.

This paper is organized as follows: Section II describes the various video recommendations existing algorithm, Section III Proposed System, section IV Advantages, section V conclusion.

A. User-User algorithm:

This is one of the oldest algorithm developed in 1992.In this algorithm the videos are suggested based on the content rating. In this algorithm the grouping is done on the basis of the neighbur hood. This could be understood more briefly with an example. Let us consider the below example that eexplains the grouping by neighbor hood. Assume that there are two users u1 and u2 watching the same video. If they are asked to give the rating u1 rated the video as 4 and the u2 as 5 for the same video. Both the users have given their ratings based on their opinions. The two users are bound in the same group because the difference distance between the two is calculated to be one. So they are grouped as neighbourhood.

The user-user algorithm seems to be fine and good at the groupings but has lot of consequences. This algorithm works fine if the ratings are given obediently and correctly based on their opinions. But few users misuse this and give their ratings casually. If this is the case then the grouping of neighbourhood would go wrong and leads to wrong analysis. Another problem with this algorithm is that there are many users across the world and they would edit the ratings accordingly. For every editing made by the users the grouping table always changes accruing to the computations and differences. So this becomes much expensive.

B. Item-Item Algorithm:

A video recommendation engine mainly suggests videos based on the view of a video and co-view of users. If a video is watched by a user, after viewing it the user will suggested to see videos. That suggested videos is mainly based on commonly watched videos. Trying to explain for suggesting videos by using firstly by co-view graph and in second figure we show the user-view graph.

By using Collaborative filtering system item-based searching is implemented. To explain co-view concept an example is taken. When a user "Z" watches X and Y videos. Based on co-view analysis those who watches video X they will also watches A,B,C videos and who watches "Y" will also watch C,D,E,F videos then "Z " user will be suggested to see A,B,C,D,E,F. Although C is repeated, only once the suggestion is given. Co-view anlaysis is shown below

CONFERENCE PAPER
Two day National Conference on Innovation and Advancement in Computing
Organized by: Department of IT, GITAM UNIVERSITY Hyderabad (A.P.) India
Schedule: 28-29 March 2014

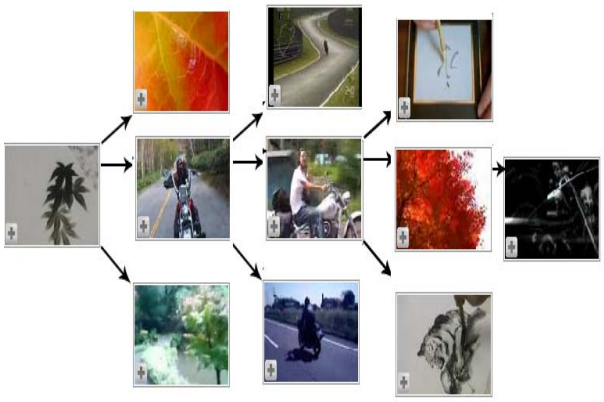

Figure: 1

Figure 1: Video-Video Co-View Graph. In this graph, every vertex is represented by a video which is linked to other videos often co-viewed. Often, only links with some minimum of views are instantiated**.**

Another anlaysis is done based on user-video graph. If a user watches a video then there will be a connection established between the user and video. Similarly all the users will be connected to the videos based on the similarity in watching videos. Suggestion of videos is made possible in above way. Suggesting is done on very small database and can take only one video of one user and analysis is made. But coming to large database computing is based on co-view statistics by making use of broadest definition if possible.

Figure 2: User-Video Graph. Another way to represent the co-view information is inherently through the user-video bipartite graph. Examining the number of paths of length 2 that exist between any two videos, gives the number of co-views.

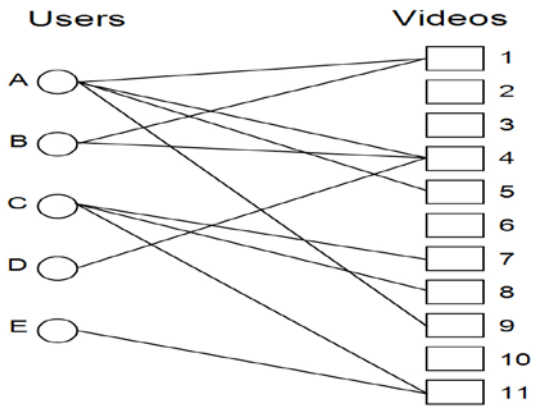

Figure: 2

C. Adsorption algorithm:

In adsorption algorithm recommendation is done by considering labels. We will call absorption to answer a question i.e.., assuming to classify a node in a graph in terms of labels to some other nodes. To answer this question based on some metrics like short distance, commute time, electrical resistance classifying is done. If two users A and B watches a common video but don't have any common interests between them. Then also the user A will be suggested to see videos watched by B because of short distance. Recommending videos is done mainly by taking into consideration multiple short paths from user node. For doing this we must take in to consideration that the video and user are not too long. When a user A with relevant to a video B when both have shortest path or several path or high degree node.

a. Absorption algorithm based on averaging:

In a graph some nodes have label. The nodes that carry some labels, forward the labels to neighbors and some collect labels that they receive. Thus each node has two roles one is forwarding and the other is collecting labels. Considering "full-information" model, let us imagine that each node keeps the history of how often it receives and at which round it received.

b. Absorption algorithm based on Random walks:

In a graph each node is assigned with a stationary probability distribution. The probability for the node is based on random walks. In the absence of absorbing node, initial choice of the node from the random walk start is completely irrelevant in determining the probability of reaching any particular node in long run. The starting point of the walk determines the probability with which the stop walk is at any of absorbing states. This implies that these probabilities are as a measure of influence of nodes on each other.

c. Adsorption algorithm via Linear Systems:

The linear system viewpoint offers natural algorithmic approach for label propagation. It also offers natural ideas for obtaining computation efficiently for all versions of the algorithm. By considering this, videos which are popular in communities are shown to a user regardless of how popular. The another benefit of the viewpoint in terms of linear system is that incremental updates to the label distributions or addition or deletion of nodes can be accommodated easily even after updating the information related to neighborhood in graph.

d. Absorption algorithm for Injection and Dummy Probabilities:

Adding a dummy label equivalent to abandoning the random walk is a very useful feature. It has advantage of slowing down random walk in a quantifiable way.

III. PROPESED SYSTEM

The proposed system is being developed to overcome the drawbacks of the existing system. In the existing system the recent updations and the notifications are not provided. The proposed system has been brought in to light to satisfy the necessities of a user which are to be considered. In this novel strategy different parameters are being developed to meet the advancements of the busy world. These updations are expected to be useful and can make many operations more efficient. This can be considered as an updated version of the existing system. This novel strategy can be explained as different modules.We have three Modules that will go deeper in to what we are trying to develop in the existing system.

Modules:

A. Module-1: Updations and notifications:

If a user is viewing few videos and later they are surfing for other similar videos, then the user is provided with all the related videos that sink to his requirements .Later even if there are any new videos uploaded then the user gets the notifications of the latest updations

CONFERENCE PAPER
Two day National Conference on Innovation and Advancement in Computing
Organized by: Department of IT, GITAM UNIVERSITY Hyderabad (A.P.) India
Schedule: 28-29 March 2014

***Example**:* If user A likes to watch a certain soap suppose X which user A frequently watches then if today the user has watched the 150th episode then the user will get the suggestion of the further updated episodes of X.

B. Module-2: Suggestions of similar kind:

If the user views videos of type A and later views the videos of type B, for the next time the user will be suggested to view the updated videos of similar type that he recently viewed. The user is also suggested with all the related videos with the parameters that are frequently viewed.

***Example**:* When a user is fond of watching videos of particular genre then in this module apart from giving the updated videos of that genre, it additionally provides suggestions related to that genre.

C. Module 3-Not Found and still Found:

If the user browses a particular video and the result happens to be "Not Found". And if this video is later updated then the video will appear in suggestions not waiting for the user to type for browsing.

Example: Suppose a user has searched for a video which does not exist and if further that video is updated then it will appear in the suggestions as soon as it is updated.

D. Advantages:

As this system has to have a purpose to be proposed, this strategy has its own advantage .the benefits behind this system are**:**

a. It reduces the amount of time spent in searching a particular video.
b. It is efficient, in a way it accomplishes half of the work by already giving a suggestion of what we often see.
c. It can be used along with popular sites also i.e this can be an additional feature in you tube.

IV. CONCLUSION

This paper has given an account of the present working strategies and the drawbacks of the various other video suggestions. This paper also gives the importance of the efficient video recommendations and suggestions to the users and overcome the existing problems. The reasonable approach to tackle the issue is the "Novel strategy". This strategy is being implemented for better and easy access to the videos according to the users preferences. The main advancement in this strategy is that the users get the notifications of the newly updated videos and also give suggestions to the user about the videos of similar kind. Which means that the user can ,to a great extent ,depend upon suggestions alone preventing the pain of new searches. This strategy is totally convenient for even the new user of that video applications. By this strategy we have chances of saving time that the system takes to upload the given text that has to be searched. Therefore there is a chance of betterment to the existing system.

V. REFERENCES

[1]. G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE-TKDE, 17(6), 2005.

[2]. B.M. Sarwarm et al., "Analysis of Recommendation Algorithms for E-Commerce," ACM Conf. Electronic Commerce, ACM Press, 2000, pp.158-167.

[3]. B.M. Sarwar et al., "Item-Based Collaborative Filtering Recommendation Algorithms," 10th Int'l World Wide Web Conference, ACM Press, 2001, pp. 285-295.

[4]. YouTube.http://www.youtube.com/results?search_query=*&search=Search.

[5]. X. Zhu. Semi-Supervised Learning with Graphs. PhD thesis, Carnegie Mellon University, 2005. PhD Thesis.

[6]. X. Zhu, G. Ghahramani, and J. Lafferty.Semi-supervised learning using Gaussian fields and harmonic functions. In ICML-20, 2003.

[7]. China Post. Youtube launches site in traditional Chinese http://www.chinapost.com.tw/.

[8]. C. Snoek and M. Worring. Multimodal video indexing: A review of the state of the art. JMTA, 25:5–35, 2005.

CONFERENCE PAPER
Two day National Conference on Innovation and Advancement in Computing
Organized by: Department of IT, GITAM UNIVERSITY Hyderabad (A.P.) India
Schedule: 28-29 March 2014