ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

Implementing Multimedia Content Aggregation for News and Music Media

Shreya Gangane, Prof. Sneha U. Bohra

Student of Master of Engineering in (CSE), Raisoni College of Engg. and Tech., Amravati, India

Assistant Professor, Dept. of (CSE), Raisoni College of Engg. and Tech., Amravati, India

ABSTRACT: Now a days, there is tremendous growth in the volume of multimedia content along with the use and growth of internet. In this paper, we are mainly dealing with News and Music Media contents that leads to the generation of large variety of content choices and each consumer has diverse preferences for content. As with the growing use of internet it will lead to the challenges for dealing with large amount of content and for the system to accurately predict what type of content to show to consumers depending on their profile. For that we propose an online multimedia content aggregation CA framework, which gathers content generated by multiple producers to fulfil its consumers demand for different types of content. Till now, most of the time multimedia content characteristics and the consumers' preferences and contexts are unknown. Therefore, there is a need to propose a self-learning algorithm for content aggregation. Our proposed algorithm is able to learn online what content to gather and how to match content and users by determining similarities between consumer types. The proposed learning algorithm guarantees both the accuracy of the predictions as well as the learning speed and efficiency of computation performed. Importantly, our system operates efficiently as the feedback from consumers and rating for the user interest is taken properly over time that is useful for support to make decisions in DSS. The preferences of consumer often depend on the consumer profile and data that helps to generate recommendation according to the categories generated for content aggregation and.

KEYWORDS: Multi-media Contents, Content Aggregation (CA), Preferences, Recommendations, personalization, Decision Support System (DSS).

I. INTRODUCTION

In recent years, with the growing use of the digitization of content leads to the use of platforms called as aggregation of content [1]. This is most probably important in the media and News industries [2]. Today with this large amount of information, aggregation of content consolidates the contents from multiple sources into one place, thereby lowering the transactions time and costs of obtaining content and introducing new information to consumers. This multimedia content aggregation platforms represent a shift away from peer-to-peer platforms operating outside of the law towards platforms that aggregate digital content within a legal framework, including Spotify for music, and Internet News for news content.

As the interest in the multimedia applications is goes on increasing and that are mostly perform with the web based platform today the new concept of personalization is growing on increasing. In our case, personalized audio/video retrieval, personalized news aggregation [3], etc. are important emerging trends which require matching multimedia content generated by distributed sources with consumers exhibiting different interests. The matching of contents are often performed by CAs [4] that also performs mining the content of numerous multimedia sources in our context these are mainly associated with music's and news in search of finding content which is interesting for the users. For identifying the content in which users are interested and performing their aggregation both the characteristics of the content and preference of the consumers are evolving over same time.

For performing the aggregation over the large amount of data that is growing on increasing and based on user's interest each user is characterized by its context, which is a real-valued vector, that provides information about the users content preferences [5]. We assume a model where users are coming sequentially with their interest to a CA, and based on the type (context) of the user, the CA requests content from either one of the multimedia sources that it is connected to or from another CA that it is the context that is used for performing aggregation uses the information parameters such as age, gender, search query, previously consumed content, etc. The CA's role is to match its user with

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

the most suitable content, which can be accomplished by requesting content from the most suitable multimedia source. After a particular content matching is made, the user is able to "consume" the content, and they can provides feedback I the form of such as *like* or *dislike which further provide the* rating. With this feedback the aggregator can help to learn the preferences of its users and the characteristics of the content that is provided by the multimedia sources.

There are two most important applications of content aggregation are present in which users are also interested and they are needy of it called as news aggregation and music aggregation. The news aggregation is mostly associated with Business news aggregators that can collect information from a variety of multinational and multi-language sources and make possible recommendations to specific individual based on their interest. On the other hand, Music aggregation is required to perform as there are lots amount of criteria's are involved like music, singer, creator and their matching listeners. So for performing aggregation with the matching listeners and matching music content they one has to select from lots amount of parameters. For distributed music aggregators we have provide the facility of sharing of music collections owned by different users without the need for centralized administrator [6].

In this paper, we have proposed the model for Content Aggregation (CA) that is performing the aggregation on two real-world datasets related to news and music aggregation. For this any CA has to collect the interest i.e. in the form of likes and dislikes of its users, both who are direct visitors to our portal or someone who is requesting from outside However, as the information of users is growing on increasing on the daily basis in the vast amount, it becomes crucial to the CA to maintain the log of interest for growing past information collected by the CAs in an efficient way, due to the vast number of contexts i.e. Parameter on which aggregation is performed and dynamically changing user and content characteristics [7]. So, for optimizing the performance of the multimedia content aggregation system, we propose an web based learning methodology that is evaluated using the notion of regret: the difference between the number of content likes minus costs of obtaining the content, that gives us the best content matching strategy given complete knowledge. In our designed framework if someone is missing to give feedback, we use indexing to denote the number of users that have visited till now.

The remaining paper is organized as, Section II, describes the needs of Content Aggregation (CA). Section III, gives the implementation of our proposed system that we have designed to perform for aggregating the contents from both music and news media sources. Section IV, shows the results of our proposed methodology that shows effectiveness of our model for performing the content aggregation. Section V gives the advantages of performing content aggregation. Finally, in section VI. We conclude the paper.

II. LITERATURE REVIEW

Cem Tekin and etal in their paper [1] suggested that a variety of content choices, consumers are exhibiting different preferences for those contents; their preferences are depends on the context in which they consume content as well as various exogenous events. For satisfying the demands of the customer for such diverse content, multimedia content aggregators (CAs) have emerged as contents from numerous multimedia sources. A key challenge for such systems is to accurately predict what type of content each of its consumers prefers in a certain context, and adapt these predictions to the evolving consumers 'preferences, contexts, and content characteristics[1].

A Hybrid Trust-based Recommender System for Online Communities of Practice Xiao-Lin Zheng and his team was conducted A case study using Stack Overflow data to test the recommender system. Important findings include: (1) learners in online cops have stronger social relations and tend to interact with a smaller group of people only; (2) the hybrid algorithm was able to provide more accurate recommendations than content-based algorithm and; (3) the proposed recommender system can facilitate the format personalized learning communities[8].

Efficient and Confidentiality-Preserving Content-Based Publish/Subscribe with Prefiltering Rapha [9] Content-based routing middleware provides multiple advantage in terms of flexibility, scalability, and simplicity for the development of distributed applications. A major obstacle to the wider adoption of these techniques is their confidentiality: messages, individual subscriptions, and containment relationships among a set of subscriptions may reveal important information about a user or group of users. While encrypted routing hides the content of messages and subscriptions, it requires costly encrypted processing algorithms that make the filtering operation computationally prohibitive for high throughput systems [9].

For making Personalization in Multimedia content search the author Emanuele Di Pascale [10], The increasing success of multimedia services poses a serious challenge to both network operators and service providers. Traditional copper-based access technologies are being replaced by fibre-to-the-X (FTTX) deployments in order to meet this increasing bandwidth demand of individual users. However, the growing traffic volume in the access segment may

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

overload the existing aggregation part of the network, therefore creating a bandwidth bottleneck in the core. To address this problem the traffic should be kept as much as possible in the access network, taking advantage of the high-capacity finer infrastructure.

III. FUNCTIONALITY NEEDS

There is large increase in the use of data and its storage today. The most famous example is the use of internet where the transaction of data is performed in the large manner. Along with this, there are large and most well-known applications that uses database in large manner, as the internet and network of network connects people to each other. As each user have their different choices in the same or different field of their choices, their preference may be same or may be different according to content. In most of the cases the preference is different as each user has different way of thinking. So server has to maintain differently each user's preference and his history information. This is the largest and most used application of the database.

As in case of News media and Music Media there are much large number of contents that one is difficult to find preferences and difficult to perform any computation on this system. There is a tremendous growth in the size and volume of content related to this both News and Music Media as the content has images, songs, videos and many more data to show to its user's. As each user have many more choices available to choose which type of data user is interested in. To show the gathered content according to user point of interest we proposed our system that is mainly related to Content aggregation (CA). It will take the users history firm the user profile and with the help search engine and prefacing finding algorithm. It initiates us, to develop a personalized search engine which delivers the content on the basis of user's preferences, profile and popularity. And a recommendation system using data mining techniques which find out users preferences and deliver content from multiple servers automatically. This leads to develop a DSS (Decision Support system) for content administrators to find out users interest to improve their services

IV. PROPOSED WORK FOR CONTENT AGGREGATION

A. Algorithms Used

For performing aggregation of multi-media content requires large amount of computation and the parameters available for this is also of various types, hence that requires algorithmic computation. The algorithms used for implementing multi-media content aggregation are describe below:

A. DISCOM algorithm: DIStributed Content Matching algorithm (DISCOM) includes a mechanism to help content aggregation. The DISCOM algorithm used in the project is as explain below:

Input: $H_1(t),H_2(t),H_3(t),T$
Initialize x into categories denoted by CA
While $t \geq 1$ do
    Run $DISCOM_{max}$ to find $p=p_i(t)$ to obtain matching action a_i
If a_i€ CA&$CA_i \neq Ø$ then Run DISCOM to obtain the contents to be selected
If a_iƐ C_i then
Show a_i to the user,
Receive feedback € 0,
If $CA_i(t) \neq Ø$ then
for j€$CA_i(t) \neq \phi$do
    send content to user
observe feedback r
end for
end if
end while

B. Pattern matching algorithm: Pattern matching algorithm will be used in searching using keywords. the pattern matching algorithm used in this module is as explain below-

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

Input keyword k, K(t), P(t), C(t) as sets of keywords, products and categories
while i>0 do
if $k_i \epsilon$ K(t) or $k_i \epsilon$ K(P) or $k_i \epsilon$ C(t) then
display marching media content $m_i \epsilon$ M(K)
end if
end while
Where,
K(t)- keywords
P(t)-products
C(t)- categories

C. Clustering algorithm: This algorithm is used to develop decision support system reports and to identify user's interest and content aggregation from multiple servers. Clustering algorithm used in this module is as explain below-

Input $c_i \not\in C$ where C is set of Categories
Initialize c_i as new category to be added into the cluster
c_i { C_i^n }
While ki ϵ C do
Select $m_i \epsilon$ M where c(m_i)= { C_i^n }
Media set ki ϵ C_i^n (M)
End while

B. Activity Diagram

Major activities during planning include understanding customer's problem, performing feasibility study, developing a recommended solution strategy, determining the acceptance criteria, and planning the development process. The product of planning is system definition and a project plan. The project plan contain the life cycle model to be used and hence we are implemented the software that works as shown in figure 1 below. The organizational structure for the project, the preliminary development schedule, the preliminary cost and resource estimate preliminary staffing requirement, tools and techniques to be used standard practice to be followed.

In the design phase the system and software design is prepared from the requirement specifications which were studied in the requirement analysis phase. The proposed system will be as shown in the below figure.

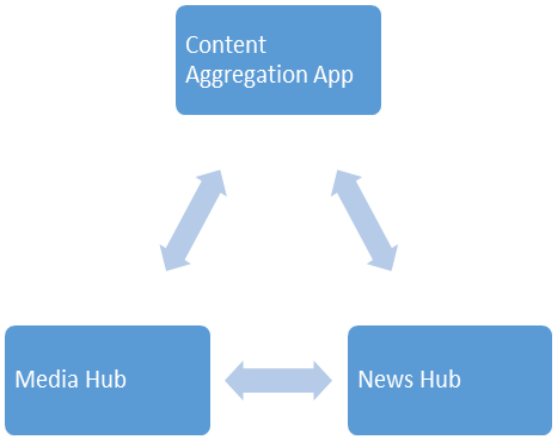

Fig. 1: Activity diagram working system

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

C. Stepwise Working of Proposed System

1. **System Initialization:** System administrator knows how system is started and configured. Although the exact detail of the source file installation is done.

2. **Media Hub-**Media hub consist of Media Category management and Content management. In the media category management there will be music and videos present. .The Content management consist of Insert /update/delete content details, View reports about Content Aggregation App

3. **News hub-**The news hub consists of two sections as News Category management and News Content management. There are various categories of news present here. The News Content management consist of Insert /update/delete content details, View reports about news content detail.

4. **End user-**the end user can perform the following task of Registration, Login/Logout, Search content, Get content recommendations, Download/read contents, Rate searched content, Like the content, Edit profile, Change password, Password recovery.

5. **Search Engine Personalization-**User can search any content using keywords. System will fetch content from two hubs as per the keyword .re-rank the content on the basis of popularity, users preferences and profile. Content delivery. Search history tracking with the help of users click through data

6. **Preference Finder-** System will automatically find out users preferences with the help of following details, Users Search History, History collected by the system while accessing the contents (users click through data), Users education profile, genders, age, profession, songs interests.

7. **Content Aggregation -** The System will collect (aggregate) the content from media hub server and news hub server depending upon the above three factors and deliver it to users as recommendation.

8. **DSS (Decision Support System) -** A decision support system (DSS) is a computer program application that analyzes business data and presents it so that users can make business decisions more easily. System will analyses historical data about searches and generate useful information about which type of users are accessing the content and which type of content are most visited, so that the administrators can enrich their contents to get more users and more popularity.

IV. CONCLUSION

Here, we try to provide, novel online learning technique for content matching by a distributed set of Content Aggregations CAs. We have characterized the relation between the user and content characteristics in terms of a relevance score, and proposed online learning system that learns to match each user with the content with the highest percentage of different categories. Our system mainly performed aggregations of the contents, that are mainly related to different categories of mainly two medias that are News hub and Music Hub. After performing successful aggregation according different categories it will helps in giving recommendations for preference finding. As the users are register to our system, our system will get the information about their age, educational background and occupations, etc. and along with this user's searches history is also helps our system to successfully provide recommendations for future searches. Our system generated the results, that helps to validate the concept of distributed content matching on real-world datasets mainly related to News and Music's and it successfully perform content aggregations and provides support for decision Making.

REFERENCES

[1] Cem Tekin and Mihaela van der Schaar, "Contextual Online Learning for Multimedia Content Aggregation", *IEEE ,TRANSACTIONS ON MULTIMEDIA*, Vol. 17, No. 4, April 2015.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

[2] Evans, D. and R. Schmalensee (2012). The antitrust analysis of multi-sided platform businesses. Coase-Sandor Working Paper Series in Law and Economics.
[3] S. Ren and M. van der Schaar, "Pricing and investment for online TV content platforms," *IEEE Trans. Multimedia*, vol. 14, no. 6, pp. 1566–1578, Dec. 2012.
[4] Eric Bruno, St´ephane Marchand-Maillet, "Multimodal Preference Aggregation for Multimedia Information Retrieval", IST-2005-2.5.10.
[5] Jelle Nelis, Dieter Verslype, Chris Develder, "Intelligent Distributed Multimedia Collection: Content Aggregation and Integration" Ghent University – IBBT.
[6] Lesley Chiou, Catherine Tucker, "CONTENT AGGREGATION BY PLATFORMS: THE CASE OF THE NEWS MEDIA", *NATIONAL BUREAU OF ECONOMIC RESEARCH*, July 2015 [Online] available: http://www.nber.org/papers/w21404
[7] The Future of Content Aggregation [Online] available: http://www.lexisnexis.co.uk/media/insights/The-Future-of-content-Aggregation.pdf
[8] Xiao-Lin Zheng, and Chao-Chao Chen, Jui-Long Hung, "A Hybrid Trust-based Recommender System for Online Communities of Practice" DOI 10.1109/TLT.2015.
[9] Barazzutti, Pascal Felber, and Hugues Mercier "Efficient and Confidentiality-Preserving Content-Based Publish/Subscribe with prefiltering Rapha" DOI 10.1109/TDSC.2015.2449831.
[10] Emanuele Di Pascale, David B. Payne, Lena Wosinska, and Marco Ruffini "Locality-Aware Peer-to-Peer Multimedia Delivery Over Next-Generation Optical Networks" J. OPT. COMMUN. NETW./VOL. 6, NO. 9/SEPTEMBER 2014.
[11] Panos Kudumakis, Xin Wang, Sergio Matone, and Mark Sandler, "MPEG-M: Multimedia Service Platform Technologie" *IEEE SIGNAL PROCESSING MAGAZINE* [159] NOVEMBER 2011.