JCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Special Issue-15, May 2019 E-ISSN: 2347-2693

Interest Based Interactivity Through Cross Platform in Big Data

Iqbal Ansari[1*], Sunil Ghimire[2], Subash Chaudhary[3], Anoop N Prasad[4]

[1,2,3,4]Department of Computer Science, East West Institute of Technology, Bengaluru, India

DOI: https://doi.org/10.26438/ijcse/v7si15.208212 | Available online at: www.ijcseonline.org

Abstract-Given the ubiquity of social media, interest-based interactivity as a main element to intensify user experience. Interest-based relevance modeling is taken out from user influence in multiple-platform social network Big Data container. The main goal of this work is to implement a platform for providing recommendation across different social network based on user interest. The streams consisted of tags from social media content through a discovery process and the application is tested on social media content streams to generate a Big Data scenario.

Keywords—*Cross platform, Data security, Big Data*

I. INTRODUCTION

A large number of social media platforms emerged in recent years with services geared towards users through adds-on such as mobile texting, Facebook with increasing popularity of Twitter, Google+ and WhatsApp, especially in entertainment contexts. These social media platforms, predominantly consisting of social networking sites (SNSs), heavily rely on individual users for content creation, in contrast to professionally produced content. With forty-one percent of the US population finding photos and videos online, interest based content discovery became the driving force for new content generation and redistribution.

As of late data have additionally turn into a surge streaming into each range of the worldwide economy. It has now swept into every business function and industry. Company's container out a large volume of transactional data, by capturing information about their operations, customers and suppliers. Also the online users, consumer devices including PCs and laptops, various new online applications, networked sensors embedded devices like social media sites, industrial machines, smart-phones and automobiles etc, have increased rapidly, thus increasing the amount of information on web. Such huge amount of data is called as Big-data. Big-data refers to a dataset, immense in nature and difficult to capture, store, manage, process and analyze with the available current technology within the good speed and time. The developing vast measure of interactive media substance has assumed an essential part in the exponential development in the substance of Big Data. This Big-data handling is a serious problem to all the companies and IT industries and also poses a heavy impact on service recommender systems. Issues like inefficiency and scalability which comes along while managing and analyzing huge amount of dataset, are of great problem for the service recommender systems.

Several social media cross-platform applications open up an area to account for single-platform content access limitations [1]. Interest-based content redistribution was facilitated by "share" function; easier content access to multiple platforms was provided through a open identity [3]; to account for increased content variety content aggregation tools, developed to combine functionalities from multiple external sources. In addition, for users, it takes time, effort and cognitive capacity to follow multiple platforms with equal dedication.

However, all these cross-platform applications continue to have limitations. Although the "share" function allowed content to be broadcasted or duplicated across various platforms, the downside of such approach was that the user could engage in one- to-many content distribution, but remained limited to receive contents from each separated platform individually [1].

Open identity facilitated access to content by allowing users to sign in to multiple websites with a single identity (ID). Such an open ID remained limited to a targeted platform rather than to multiple parallel platforms. Content aggregation platforms in turn provides users with larger amounts of content access, yet does not support interaction and content discovery through other user experiences.

To account for the above mentioned limitations, design a unified access model to interest-based content modelling. Capitalized on existing SNSs to create a Big Data repository – term used to describe a large and complex collection of growing dataset that is difficult to manage and process using traditional database management tools – to model an interest-based content segmentation and content discovery through user interaction.

II. Related Work

H. M. [2] Inc. (2013) Social media management. [Online]. Available: https://hootsuite.com/Single platform: It uses “share” function to send interest based content data to multiple platforms through open identity[3] which uses combine functionality from multiple sources. But the disadvantage is it takes more time to effort combining and it takes time to check multiple platforms takes place.

C. J. Jacoby [3]Nowadays, keyword inquiry keeps on being basic, if not the most widely recognized, strategy used to scan revelation record accumulations for conceivably applicable material. Unfortunately, keyword inquiry offers just a defective answer for this issue on the grounds that it misses numerous important records. By its basic nature, keyword search is sharply engaged to locate the definite terms determined in the query. Search for documents containing the word “car” and you’ll find exactly what you’ve requested. However, you’ll miss documents containing potentially relevant words like “automobile”, “Ford,” “GM,” and “Toyota.” One can compensate by increasing thenumber of search terms and by adding stem searching to find plurals of words, but the results will still likely overlook some relevant materials.

Algorithm for keyword match

Step1: Input ‘n’ number of string and keyword

Step2: for (i=0; i++; i<n) {

for(j=0 ; j++; j<m && i+j < n)

if(n of words[i+j]! = keyword[j]) break; // confound found, break the inner loop

if(j==m) // match discovered

}

G. Adomavicius, and A. Tuzhilin[05] A survey also describe that the traditional recommendation approach are based on rating and rankings only that is content based without considering the interests and choice on an user. The recommendation provided did not considered the personalized choice of the users. It also describes various limitations such as scalability and transaction time that are faced by the current service recommendation methods and certain extensions that can make better recommendation capabilities and make recommendation systems appropriate to a broader range of applications.

Z. D. Zhao, and M. S. Shang [4] Tells about various problems faced by Collaborative Filtering algorithm and how to overcome those problems, a widely used personalized recommendation approach, is its scalability, i.e. when the size of the dataset is very huge, the processing expense of Collaborative Filtering would be huge. The issue of substantial scale reckoning assignment can be overcome by using the cloud computing platform. Implement the Collaborative Filtering method on the cloud-computing platform to take care of the adaptability issue of recommender system, Hadoop, which solve the scalability problem for large scale data by dividing the dataset.

Algorithm for collaborative filtering

Step1: for each url in url list catalog , u1

Step2: for each user C who rated , u1

Step3: for each url u2, rated by user C

Step4: Record that a customer rated u1 and u2

Step5: For each url u2

Compute the similarity between u1 and u2.

E. P. Bucy [6] Tells about cross platform applications for users, it takes time, effort and cognitive capacity to follow multiple platforms with equal dedication. However, all these cross-platform applications continue to have limitations. Although the "share" function allowed content to be broadcast or duplicated across various platforms, the downside of such approach was that the user could engage in one-to-many content distribution, but remained limited to receive contents from each separated platform individually.

O. Foundation [7] (2013, Sep 13) Openid foundation website[Online]*http://openid.net/*facilitated access to content by allowing users to sign in to multiple websites with a single identity (ID). Such an open ID remained limited to a targeted platform rather than to multiple parallel platforms. Content aggregation platforms in turn provided users with larger amounts of content access, yet did not support interaction and content discovery through other user experiences. Above mentioned limitations can be overcome by using a unified access model to interest-based content modelling.

S. J. McMillan[8] To implement crossplatform geared towards usercentric experience, utilized the following six dimensions of perceived interactivity. Direction of communication– the ability to receive return messages, in addition to one-way communication, prevalent in traditional mass media such as TV or radio; time flexibility – the ways to engage in communication in real time as well as being able to retrieve archival conversations; sense of place – creation of a common virtual communicative context; level of control – user agency to select responses in the most appropriate context, ability to track messages and identify if they were delivered; responsiveness – relates to the concept of sequential interrelation between messages where three-message-sequence model was proposed as an ideal one and perceived purpose of communication – which identifies goals of communication and makes interaction meaningful.

E. J. Downes and S. J. McMillan [9]To implement six dimensions, focused on all three proposed interactivity traditions such as, Human-to-human interaction – the interaction between users in a mediated environment. Human-to-document interaction – user interactions with a specific content, created by other users and with the creators of the content themselves. Human-to-system – adaptive system that enables users to get higher levels of control and encompasses the previous two dimensions by combining interface components and human components. Human-to-system interaction includes the methodologies such as subscription management, content management and content interactivity. Human-to-human interaction which helps through rating mechanisms users indirectly inform other users about the relevance of a certain item to ensure coverage of important issues.

P. Wilson[10] Identified seven techniques to ease the information overload associated with interactive Big Data models: information retrieval, aiming at finding information pertinent to a given subject through the use of keywords ; information filtering, relying on filtering techniques to highlight relevance from a continuous flow of information ; rank filtering, providing omission techniques to identify relevant items, using predefined factors such as the number of recommendations, user acceptance and popularity within the community; brute-force interaction, defining techniques that enable immediate and effortless initiation of interaction; content approximation, to help users selecting the most important and relevant items by providing users with a brief preview of a given item extracted from each of the properties; contextualization, introducing techniques to organize the information, i.e. to highlight its significance; and information stack, combining the aforementioned techniques with actions to postpone and redefine the priority of an item.

Daniele Dell[12] present the concept of Semantic Web-enabled Recommender System, based on the retrieval from the linked data Web of the necessary pieces of knowledge about items and users. We illustrate the general structure of this new family of Knowledge-based Recommender Systems and we explain how we concretely followed this approach to develop a tool to recommend Web services in the context of the SOA4All project. We also offer our considerations about the strengths, the current limitations and the possible extensions of our proposal.

III. METHODOLOGY

A. SYSTEM ARCHITECTURE

System architecture is the reasonable configuration that defines characterizes the structure and conduct of system. In the system architecture shown in Figure 3.1, Content/Subscription management of user is done in the registration phase. Then the interests are extracted from different websites like youtube, facebook, twitter by using the web API's such as youtube data API protocol, Public content solution API, Twitter JSON API respectively and stored in the information extraction repository and then User browsing behaviour is extracted and stored in the interest mining repository. Lastly content matching can be done between the information extraction and the interest mining repository and finally content recommendation is given to the user.

User Management: Here the User registers to the system. Manage content filtering and browse and view content using this module.
Also this module implements the blacklisting of contents not matching to user interest.

Information Extraction: The contents are extracted from different websites like youtube, facebook, twitter are stored in Information Extraction Repository.

Interest Mining: Based on user browsing behaviour on contents, this module learns the user interest and constructs user profiles grouping user of similar interest.

Content Matching: This module will match the contents to user interest based on meta data matching and also collaborative recommendation and provides content recommendation to the user.
The entire system will run on Hadoop Cluster.

The system architecture is shown below.

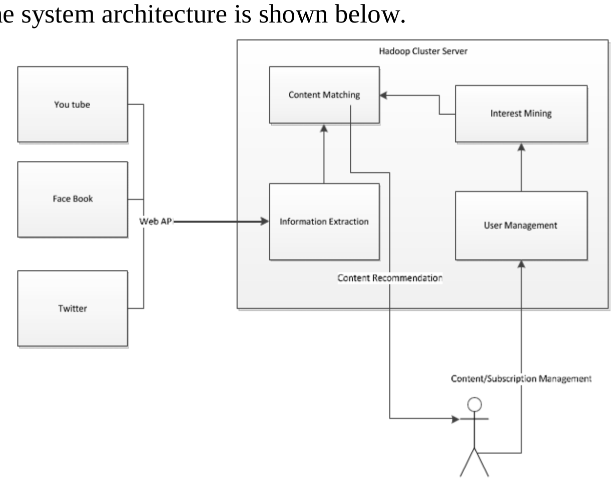

Figure 3.1 System Architecture

B. DATA FLOW DIAGRAM

Pictorial description of the moving of information through an information system is called as data-flow diagram.

a) Level 0 Data flow diagram

A setting level or level 0 information stream graph demonstrates the communication between the outer operators and framework which go about as information sinks and information source. On the situation of the framework's connections with the outer world are demonstrated effectively regarding information transmission streams over the framework limit. The graphical representation gives no piece of evidence to its internal organization and shows the entire system as a unit process.

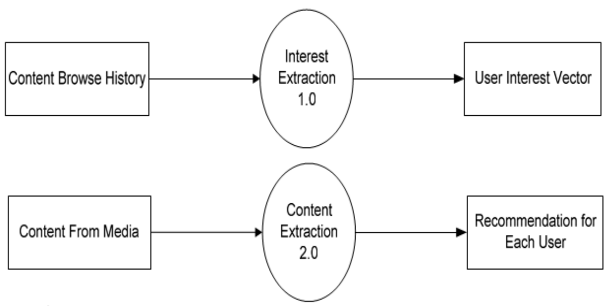

Figure 3.2: Level 0 Interest and Content extraction Data Flow Diagram

A Figure 3.2 explains Level 0 the content browse history and content from media were the two systems. Building user interest vector and recommendation for each user were the two external agents. Interest and content extraction provides the interaction between the system and the external agents.

b) Level 1 Data flow diagram

The Level 1 data stream graph speaks to how the framework is isolated into sub- framework, each of which manages information transmission streams to or from an outside operator, which together give each operations of the framework as single. It demonstrates inward information and demonstrates the stream of data between the diverse parts of the framework.

Figure 3.3 explains Level 1 system content browse history is divided into subsystems such as extract content metadata, build feature vector, user interest vector which together provide every functions of the system as a whole.

System content from media is divided into subsystems such as extract content metadata, construct feature vector, matching to interest vector, ranking the content which together provide every operations of the system as a single and finally provide the recommendation to the user.

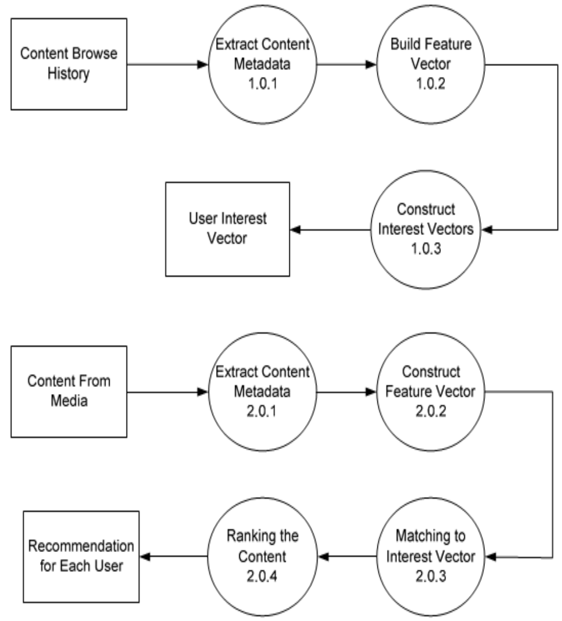

Figure 3.3: Level 1 Interest and Content extraction Data Flow Diagram

IV. CONCLUSION AND FUTURE SCOPE

Conclusion

A platform is designed which encompasses theinterconnectedness between services and social media platforms. It Highlights user value and bridge user needs through social media and professional content by enabling users to model interest-based relevant streams in Big Data.Proposed approach extended the work of previous implementations of user centric services of Big Data that were predominantly geared to strengthen internal services through multiplatform content access and fluid content sharing. This study is based on the interest-based architecture that leads to access professional and social media content streams posted to the radios. Adaptive Big Data User-centric model capitalized on a flexible environment, sensitive to changing data fluxes between services provided by social networking sites.

To overcome limited content access, platform interoperability issues, and lack of content segmentation across multiple platforms. The interest will be corrected and an interest gets enhanced.

Future Work

Furthermore, even proposed potential solutions to account for content overload and interactivity, future studies should test social shaping of this application by the users. Future studies should also address potential shifts in business models for media companies that would go in accordance with an expanded interest-based goal of Big Data User-centric model.

ACKNOWLEDGMENT

Firstly, we express our sincere thanks to our guide Mr Anoop N Prasad Assistant Professor, Department of CSE, EWIT and Dr.ArunBiradar, Head of Department,Computer Science and Engineering for their moral support. We express our sincere gratitude to our principle Dr. K Chennakeshavalu for his constant support and encouragement; we also thank all the faculties of East West Institute of Technology for their co-operation and support.

REFERENCES

[1] AstaZelenkauskaite and Bruno Simoes, "Big data through cross-platform interest based interactivity"Big Data and Smart Computing, vol. 47, no. 2, pp. 191-196,2014.

[2] H. M. Inc. (2013) Social media management. [Online]. Available:

[3] C. J. Jacoby, "Understanding the limitations of keyword search," Equivio, white paper, 2012.

[4] Z. D. Zhao, and M. S. Shang, "User-Based Collaborative-Filtering Recommendation Algorithms on Hadoop,"In the third International Workshop on Knowledge Discovery and Data Mining, pp. 478-481, 2010.

[5] G.Adomavicius, and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of- the-Art and Possible Extensions," IEEE Transactions on Knowledge and Data Engineering, Vol.17, No.6 pp. 734-749, 2005

[6] https://hootsuite.com/E. P. Bucy, "Interactivity in society: Locating an elusive concept," The information society, vol. 20, no. 5, pp. 373–383, 2004.

[7] O. Foundation. (2013, Sep 13) Openid foundation website.[Online]. Available:http://openid.net/.

[8] S. J. McMillan, "A four-part model of cyber-interactivity some cyberplaces are more interactive than others," *New Media & Society*, vol. 4, no. 2, pp. 271–291, 2002.

[9] E. J. Downes and S. J. McMillan, "Defining interactivity a qualitative identification of key dimensions," *New Media & Society*, vol. 2, no. 2, pp. 157–179, 2000.

[10] P. Wilson, "Interdisciplinary research and information overload." *Library Trends*, vol. 45, no. 2, pp. 192–203, 1996.

[11] Mohamed sarwat, Justin J. Levandoski , Ahmed Eldawy , and Mohamed F.Mokbel," LARS*: An efficient and scalable Location –Aware Recommender system.

Daniele Dell'Aglio, Irene Celino, and Dario Cerizza, "Anatomy of a Semantic Web- enabled Knowledge-based Recommender System

