The Multi-media Content Aggregation and Preference Finding System: A Review

Ms. Shreya R. Gangane, Prof. SnehaU. Bohra

Abstract— This paper gives a brief review for the multimedia data that is focused on the proper collection of personalized multimedia contents based on the concept of Content-Aggregation (CA). As there is increase in the amount of multimedia data by the use of internet,our system aims to aggregate continuously the available multimedia contents in its providers and suggest efficiently to its end-users. Thus, the system will allow users to receive multimedia content that is automatically and specifically composed of the reach amount of information. Here, is the need to perform the specific aggregationof the contents based on the user's content profiles based on dynamic context parameters. By performing the literature study, all the proposed mechanisms for content aggregation are articulated in a flexible and robust architecture. There is the need for theproper aggregation mechanism to get useful content at proper time from large amount of multimedia information/data available with us. Here, we have given the brief description of the terms that properly satisfy the functionalities of the multimedia Contents and their Aggregation. We also studied some of the algorithms and techniques that are useful for multimedia content aggregation.

Index Terms— Multi-media Contents, Content Aggregation (CA), Preferences, Pattern-matching, Multi-model Data Cube.

1 INTRODUCTION

TODAY, multimedia is everywhere and it becomes as the daily need to come up in our way. This happens mainly due to the growing the demand and use of the internet [1]. The devices that are attached to the Network are become commonplace and the Internet services enablingusers for sharing or remote viewing purposes are growing on increasing. There is a tendency todistribute a personal multimedia collection, due to a lackof organizational skills, limited storage capacity or functionality demand. This means that several usability problems are faced by the end users as they unable to fullybenefit from their own multimedia collection. A multimedia content aggregation is difficult to organizeand easily resorts in chaos, such that users are unknown about where they stored their desired contents.

In recent years, the digitization of content has growing on increasingon platforms of aggregation of contents in many important industries as well as in themedia and Internet-based industries. These new platforms of Content Aggregation consolidate content from multiple different sources, and help tolower the transactions costs of obtaining content. Aggregation is the process in which two or morelinguistic structures are merged to form a singlesentence. It helps in generating concise and fluent text [2]. Such aggregation platforms represents a shift away from platformsoperating outside of the law and towards the platforms that aggregate digital content. This includes the areas asSpotify for music, Hulu for movies, and Google News for news content means broadly speaking in the Entertainment, social and news media.While an extensive literature has focused on pricing by platforms, little is knownabout how the quantity and quality of content provided by platforms that operate within legal bounds in essences consumer search and the resulting consequences for contentproviders.

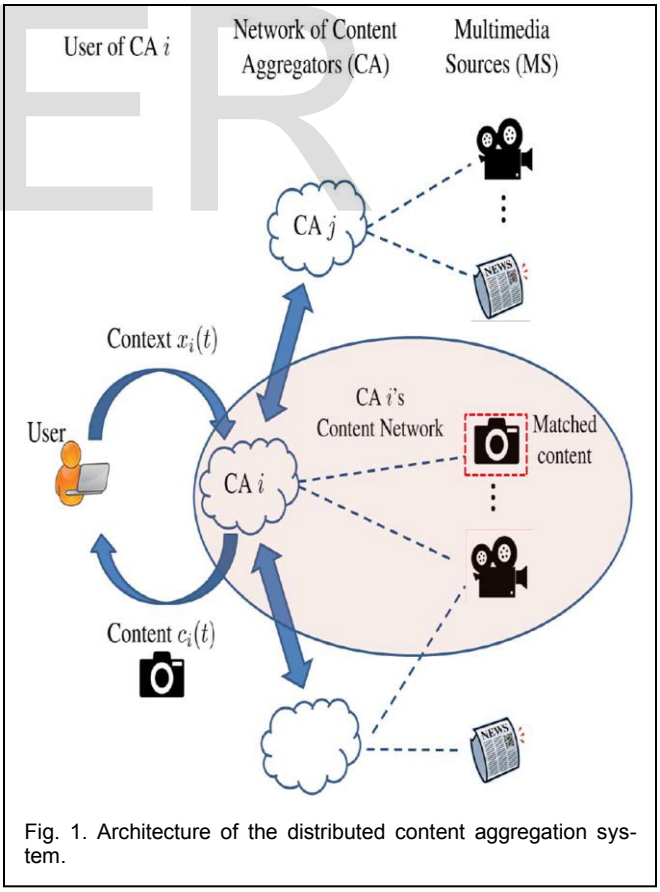

Fig. 1. Architecture of the distributed content aggregation system.

The multimedia applications including web-based TV, personalized video retrieval [4], personalizednews aggregation, etc. are growing their application areas that requires matching multime-

dia content generated by distributed sourceswith users exhibiting different interests. The content matching isperformed by Content-Aggregation (CA). The CA's role is to match its user with themost suitable content, accomplished by requestingcontent having suitable multimedia sources. The use of feedback that helps a CA to make the decisionfrom the user's point of view and the characteristics of the multimedia contentsources. This also performs the content learning and as there is mostly the use of internet it is termed as online contentlearning and the aggregation performed and forms the basis of the Design Support System (DSS). With the growing demand, there are two possible real-world applications of content aggregation as business news aggregation and music aggregation [3].An example architecture of the system with its users, CA of different multimedia sources are given in Fig. 1 [1].

More importantly, it offers the multimedia collection in such a way that any user can consume their multimedia. Not necessarily having any technical knowledge. When a multimedia items are not supported by a particular device due to a code incompatibility, the system takes care of this. And when the contents are incompatible that will not be shown. The paper is organized as follows, Section II discusses the literature review that makes the study of some author who have tried to solving the problems put forward in this section. In Section III, some important terms are defined along with their functionality needs that helps to understand our topic more precisely. Finally, in the Section IV, gives some algorithms and techniques to perform multi-media content aggregation.

2 Literature Review

Representing and fusing multimedia information to discover semantics in multimedia and determining semantic concepts by iteratively and interactively refine their queries are the key issues in multimedia content-based retrieval. More specifically in this paper, authors address the problem of multimedia content retrieval through both of the original multimodal information representation and also of a machine learning-based fusion algorithm. Here, author [5], first define a novel preference-based representation adapted to retrieval problem, and then, investigate the Rank Boost algorithm to combine those preferences to fulfill a user's query. This system ends up with being a flexible retrieval model and is blind to the intrinsic properties of the multimodal information that is fed as an input. This approach is tested on annotatedimages and compared with his second strategy of SVM-based fusion. The resultsof this show that our approach is same in performance point of view but,contrary to SVM is that, it is parameter free and faster. The Relevance Feedback loop obtained here allows users to build complex queries that are made out of documents marked as positive and negative examples. From training set, a learning process creates a model of concept from a set of data features that is finally important to provide relevant documents to the user.

In this paper [6], author proposes a new layer for a novel IPTV service that focused on distribution of content that is of personalized multimedia format over IP networks based on the concept of content-zapping, in contrast to traditional channel-zapping. This system aims to aggregate available multimedia contents in its providers. This system also suggest efficiently its endusersthat which contents are the mainly advantageous. Thus, the system allows its users to receive a multimedia streaming that is automatically composed for them. At the same time, the system will provide services of dynamically classifying its users, multimedia contents,and multimedia providers. This classification dynamically built and gradually refined based on the interactions between all the entities of that system. Then, the specific content are suggested to one or more users which is the most essential thing. This contents are mainly based on the user's content and user's profiles as well as based on dynamic context parameters, such as the user location, current date/time, the client device, the user's environment, etc. For synthesizing and saving all the knowledge user have, a classification methodology, learning algorithms and recommendation techniques were defined. All mechanisms are in a flexible, efficient and robust architecture. A client prototype was developed and tested with the proposed system.

In case of social media, multimedia content is spread around the home network of people and content services on the Internet, with different service providers such asYouTube, Flickr, and Facebook. In this paper [7],authors present a systemthat performs the aggregates of all the multimedia content and then integrates it into a unified collection according to the user's convenience.
The CA system provides transparency of locationfor the multimedia content,content filtering on player compatibility to aid in improved usability. This enables userto rediscover his multimedia collection not having technicalknowledge. Here, author performs, a proof-of-concept implementation called as Intelligent DistributedMultimedia Collec-

tion (IDMC) that has been made which is able to detect and browse UPnP mediaServer devices aswell as collect information from YouTube. This implementation consist of a media player and is able to control UPnPMediaRenderer devices remotely.
In this paper,author presented an aggregatedview of all the content along with the contents storedin the cloud. This can also be achieved by allowing the pluggable dataproviders for adding content to the system. When multimedia content gets added to the system, it get integrated in the unified collection.
This pipeline system removes duplicates, merges the metadataof the item being added thisperforms metadata completion with external information sources.Here, some kind of batching technique should be available to be usedsince adding content items one at a time becomes slow when reaching large collections that reaches over 1000 items.

With the today's digital world and increasing use of the news media, the digitization of content has led to the starting of the platforms that draw information from multiple heterogeneous sources [8]. Policymakers are concerned that these new platforms are threatening incentives for the productionof original content. As a result of this, policymakers would force aggregationplatforms. This is required for paying or require an explicit "opt-in" for content providers. For understanding the possible consequencesand underlying rationale of such laws, Authorexplore whether aggregation of content by using a single platformthat encourages users to "skim" content or to investigate in depth. Author study a contract dispute that ledto an aggregator that removes information from a major content provider. Author find out that after the removal, users were less likely investigate additional, related content in depth.

3 Some Important Terms and their Functionalities

3.1 Multimedia Contents

Now a days, multimedia has become an inevitable part of any format of presentation. It has variety of applications right from entertainment, news, education and many other. The increasing use of internet give rise to the evolution of the multimedia content.
Multimedia can be defined as, "Multimedia isthe media that having multiple forms of information content and information processing like text, images, audio, video, animation, etc. toinform or entertain the user [9]. *In another way, Multimedia can*also be refers to the use of electronic media to storeand experience different multimedia content in multimedia format.
Multimedia can be said as the traditional mixed media infine art, with a broader scope withthe "rich media" for interactivemultimedia contents. Elements of this System and the computer information in the multimedia formatcan berepresented through different format as audio,image, video and animation along with traditional media like text and graphics.Multimedia system finds its application in various areas are not limited to but including many sectors of advertisements, education, entertainment, engineering, medicine, mathematics, business, scientific research and spatial, temporal applications, and lot much. Hypermedia can also be considered as one type of particular multimedia application.

3.2 Content/Data Aggregation

Content aggregation is any process in which large amount of information is gathered and expressed in a summary form, for many different purposes such as statistical analysis, recommendation system [10] .In the common aggregation process the purpose is to get more information about particular groups based on specific attributes such as age, profession, income, etc. The information that can be collected about such groups is used for Web site personalization for choosing the content and advertising that likely to appeal to individual who is belonging to one or more groups for which data has been collected. This can be very much useful for online analytical activities. For example, a site that sells music CDs that might make some advertisement of certain CDs based on the age of the user and the data aggregate for their age group. One of the most common type of content aggregation process id Online analytic processing (OLAP). The OLAP is a simple type of data aggregation, which is used by the marketer as an online reporting mechanism to process the information.Data aggregation can be personalized for personal data aggregation services that offer the user to collect their personal information at the single point from other Web sites. The real life example of this is as, the customer uses its private single master personal identification number (PIN) for accessing them to various accounts for financial institutions, airlines, book and music clubs, and many other also.

3.3 Characteristics of Content Aggregation

1. *Timeliness* -Aggregators should update information on daily basis with feeds that may not be visible on a web platform if behind a paywall.

2. *Accessibility* –The Aggregation system should use the concept of indexing that guarantees your search will be more precise and thorough. The aggregated content should be accessible easily by personal users easily.

3. *Usability* –The aggregated contents are usable in many different application in an online basis. A single point of access through an aggregator is your gateway to an uninterrupted supply of information for business continuity.

4. *Indexing* –The aggregated contents uses the Indexing method the offers searching in precise and comprehensive way, ensuring you reach the information you need faster. Filtered, targeted, styled and formatted and results obtained by it needs minimum processing.

3.4 Functionality Needs of Content Aggregation

1. The term Aggregation must focus on meeting end user needs [11]. It is no longer enough to have information if it is not available to a user at the appropriate time. The ongoing development of methods like indexing and search is a necessary requirement in managing content overload. The Aggregationprofessionals use indexing for both precision, and for surveying the breadth of information. The information-push tools or product environments where end-users access information require these capabilities, these are often working behind the scenes.

2. Aggregation method need to retain supply of content and this needs to be achieved through continuing to offer publishers a valued distribution channel. This can also increasingly, becoming a component of publisher digital strategies.

3. Aggregation will increasingly need to address the challenge of social and online media and News Media hubs. These aggregators have fully incorporated above media types in analytics, media and sentiment measurement tools, the challenge of utilizing these rich sources of information in conjunction with licensed content has not been fully utilized. The challenges are mainly due to the stringent licensing, copyright and quality controls imposed by their business models. But at the same time this does not, however, prevent them from securing and distributing trusted online content.

4. Aggregation need to continue to invest in the full value chain for actively differentiate themselves from free or low-value services. It should act as the combination of high-value content and the associated services in making it easier to interrogate and extract information which should bring customers to services.

5. An another important job of the aggregators is to acknowledge specific job functions or key segments by building products or solutions that deliver relevant content. This is where information can be applied to a specific task with a defined value. This makes it easier to justify the costs of services by using aggregated contents.

6. The most important need of the content aggregation is that, the attributes of information on which decisions are based thatshould come from authoritative, reviewed and edited sources - many of which are behind publisher or aggregator paywalls.

4 Content Aggregation Algorithms and Techniques

4.1 Aggregation by Preference Finding Algorithm

In situations such as conference program committees andfaculty searches, a group of people are have to review a large set of candidates, and then they have to collaboratively identify which of these candidates are best. This problem is from the domain ofgroup recommender systems [12]. Whilemost recommender systems use the preferences of a largegroup of people to extrapolate those of an individual, grouprecommender systems aggregates individuals' preferences torecommend the best items for the group as a whole.In particular,these challenges include the need to accurately and fairlyaggregate users' individual preferences, and the difficulty ofdesigning a user interface that facilitates collaborative decision making as the Decision Support System.

The preference aggregation algorithm is fit forsitu-

ations in which a limited number of users that reviews asmall subset of a large set of candidates. Thisalgorithm aggregates scores by using relative preferences of users that is making the search for a Kemeny-optimal ordering of items. One canuse variable-neighborhood local search, allowing to the fast and relevant discovery of high-quality consensus orderings, which facilitates effective categorization of candidates while being feasible for large problem instances. Thisalgorithm provides a significant increase in solution quality. The potential applications of thisalgorithm is in group recommender systems for a variety of scenarios that includes the program committees and faculty searches,and a number of considerations regarding its use.

4.2 Content Aggregation by Pattern Matching Technique

Consider the environment where a subscription system is continuously evaluating pattern-based requests over sequential data having no specified bounds. Here for the aggregation, an extension of the traditional pattern-matching techniques that is for efficiently handling such continuous queries is useful technique. This extension introduce the variables in order to augment their expressivity in some particular pattern. On the basis of the extended class of parameterized queries, our main contributions are suggested in the following way. Firstly, we define a refinement relation based on variable relaxation. Secondly, we use the semi-lattice structure of parameterized patterns for patterns aggregation and filtering. Here, we have to use an on-line pattern aggregation algorithm for reducing the cost of pattern-matching and for filter out sequences that doesn't not match any of the patterns in a subscription cluster [13]. Finally, we can show thethrough analysis and experiments, that our techniques reduces the cost of the matching process from the large set of data available with us.

4.3 Multi-dimensional Content Set Aggregation as Data Cube

Most of all the multimedia data can be viewed or shown in the multi-dimensional format. Data analysis applications tries to aggregate the data fromlarge dimensions that is looking for anomalies or unusual patterns [14]. The SQL aggregate functions and the GROUP BY operator produce different dimensional aggregates as zero-dimensional. Some applications can also need the N-dimensional generalization of these operators for different application requirement. Here the aggregation operator is called the data cube or simply cube. The cube operator generalizes in the form as the histogram, cross-tabulation and different techniques like roll-up, drill-down, and sub-total constructs found in most report writers and also helps in the OLAP operation. The novelty is that cubes are relations, and the cube operator can be imbedded in more complex way and non-procedural for the data analysis programs. The cube operator treats the N aggregation attributes as a dimension of the N-space. The aggregation for the set of attribute values is a point in this space. The Super-aggregates are then computed by aggregating the N-cube to lower dimensional spaces. These technique includes the cube and roll-up, drill-down, like operators, how they fit in SQL, defining of new aggregate functions for cubes, and efficient techniques to compute the cube. Many of these features are useful to perform the content set aggregation as data cube.

5 Conclusion

In this paper, we studied the technique of the multimedia content aggregation which is the novel techniquemainly used for content matching from the distributed set of multimedia contents. We have seen that with the growing use of internet the data from the multimedia content sources and data is growing on increasing. This multi-media data is mostly from the audio, video and news sources. There are number of technique that we have learned in the literature that works on the method of multimedia content aggregation.The user and content characteristics differentiate the technique used for the Content Aggregation (CA's). In the further section, we have studied the characteristics and the definition of the Multimedia, Data Aggregation and the features and needs of Content Aggregation (CA). As we are doing the content aggregation it needs the some algorithms and techniques, for this we have perform some manipulation on the large amount of distributed and heterogeneous data. Here, we have learned about the preference finding, pattern matching algorithms and the technique for the content set aggregation on the multidimensional data that is stored on the data cube. The content aggregation (CA) method and preference finding system for multimedia data which we have studied, is very much useful for viewing the important information from the large amount of the collected contents.

REFERENCES

[1] 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.

[2] Reiter, Ehud and Robert Dale. 2000, "Building Natural Language Generation Systems", Cambridge University Press, Cambridge.

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

[4] 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.

[5] Eric Bruno, St´ephaneMarchand-Maillet, "Multimodal Preference Aggregation for Multimedia Information Retrieval", IST-2005-2.5.10.

[6] Rodrigues, J. ; Salvador, P. ; Nogueira, A., "Multimedia content aggregator applied to an IPTV content-zapping service", EUROCON - International Conference on Computer as a Tool (EUROCON), 2011 IEEE.

[7] JelleNelis, Dieter Verslype, Chris Develder, "Intelligent Distributed Multimedia Collection: Content Aggregation and Integration" Ghent University – IBBT.

[8] Lesley Chiou, Catherine Tucker, "CONTENT AGGREGATION BY PLATFORMS: THE CASE OF THE NEWS MEDIA", National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138. July 2015. [Online] available: http://www.nber.org/papers/w21404

[9] Unlocking the Power of Multimedia Content [Online] available: http://promotions.prnewswire.com/rs/prnewswire/images/ebook_multimedia-content.pdf

[10] Data Aggregation Definition [Online] available: http://searchsqlserver.techtarget.com/definition/data-aggregation

[11] The Future of Content Aggregation [Online] available: http://www.lexisnexis.co.uk/media/insights/The-Future-of-content-Aggregation.pdf

[12] Preference Aggregation in Group Recommender Systemsfor Committee Decision-Making[Online] available:ftp://ftp.cs.brown.edu/pub/techreports/09/cs09-07.pdf

[13] du Mouza, C. , Rigaux, P. , Scholl, M., "On-line Aggregation and Filtering of Pattern-based Queries", 18th International Conference on, Scientific and Statistical Database Management, 2006.

[14] Jiawei Han and MichelineKamber, "Data Mining Concepts and Techniques" (2nd Edition).