ICCCM-JOURNAL OF SOCIAL SCIENCES AND HUMANITIES
2023; 2(4): 1-8
Published online 07 24 2023 (https://icccmjssh.com/)
doi: https://doi.org/10.53797/icccmjssh.v2i4.1.2023
e-ISSN 2811-4469

Big Data Push Strategy of Media Convergence Advertising Communication

Xu, W.[1] & Ariffin, A[2]

[1]School of Cultural Communication, Liming Vocational University, Quan Zhou,362000, CHINA
[2]Faculty of Technical and Vocational Education, University Tun Hussein Onn Malaysia,Batu Pahat, 86400, MALAYSIA

*Corresponding Author: hasnida@uthm.edu.my

To Cite This Article: https://doi.org/10.53797/icccmjssh.v2i4.1.2023

Received 24 June 2023, Revised 8 July 2023, Accepted 22 July 2023, Available online 24 July 2023

Abstract: The key to the accurate match between advertisement and target users is to accurately analyze the user's interest points,and then to put the advertisement which is consistent with the user's interest. Data Management Platform is not only a user data analysis platform,but also the core of the whole network advertising push service system. Its main function is to analyze the characteristics of users' online interest. This paper studies about the big data push strategy of media convergence advertising communication. Firstly,study the web page feature extraction technology,similar web page aggregation method,user behavior analysis algorithm and other user behavior feature analysis methods. Secondly,this paper proposes an optimization strategy based on Redis to improve the accuracy of similar web page aggregation. Finally,it is combined with the original vector space model and clustering algorithm used in media advertising aggregation to redefine the aggregation standard,so as to improve the accuracy of advertising push aggregation. The results of the experimental data show that the above methods are effective, and the analysis of user interests is more accurate, thereby accurately matching integrated media advertisements that match user interests, making the big data push strategy of integrated media advertising dissemination more effective.

Keywords: Target Users, Matching, Data Management Platform, Media convergence

1. Introduction

Due to the rapid development of the internet and the continuous expansion of internet users,internet shopping has become more and more popular. In this context,internet advertising came into being. However,with the advent of the era of big data,the extensive and aimless disadvantages of traditional internet advertising are becoming more and more obvious,which makes the market urgently need a kind of accurate internet advertising,so in this context,a new advertising mode,network oriented advertising in the big data environment,is born,that is,through certain technology to push advertising to target users (Jiang Enlai,2009; Deng Xinmin,2006) . Directional advertising with strong market demand and its own advantages will usher in great opportunities for development. The academic research on this new advertising mode and the industrial practice are still in the stage of development (Ma Deyong & Sun Mengxin,2014). Under this background,it has certain theoretical significance to carry out the research on network directional advertising technology(Chen Hong & Liang Junmin,2013). It can not only greatly develop the advertising technology,but also create considerable economic benefits and create a win-win situation. Online directional advertising is the core of online advertising push service based on big data,so it has a certain practical significance to carry out the research on it in time.

"Media convergence" is a modern propaganda strategy that aims to utilize various media platforms such as radio, television, and newspapers, by integrating their similarities and complementarities in terms of manpower, content, and publicity. This strategy involves accommodating resources, content, publicity, and interests to fully utilize media carriers. (Liu Haiye,2022).

2. Network oriented advertising push service system

This section first introduces the functional architecture and specific process of network directional advertising push service,and then from the two parts of Data Management Platform and demand side platform,expounds some main

**Corresponding author: hasnida@uthm.edu.my*

functional modules in the two platforms,such as user behavior characteristics analysis,Real-Time Bidding (RTB) and so on(Wang kun kun,2014;Sun Shaoyi,2008;Yang Xiaohui,2009). Finally,the functions of these functional modules and the realization principle of these functions are explained in detail.

2.1 Definition of media convergence

"Media convergence" is first of all an idea. This idea takes the development as the premise and the promotion of excellence as the means to bring the advantages of traditional media and new media into full play,so as to turn the competitiveness of single media into the common competitiveness of multimedia,so as to serve "me" and "me". "Media convergence" is not an independent entity media,but an operation mode that integrates and utilizes the advantages of radio,television and internet,so as to comprehensively enhance their functions,means and values(Xu Xiang,2022). It is a practical scientific method and a concrete behavior that can be seen and felt in the practice.

"Media convergence" is a unit,a voice and a price. Radio,television and Internet become three forms,means and methods of serving a project activity at the same time. The price will be much higher than any single media,and the customer's recognition of this activity will be greatly improved. In the past,we used to rearrange the advertisements in TV programs when we went online and take out the advertisements. Now,when many customers come to talk about advertising,they take the initiative to negotiate with the "Media convergence" price,and they also keep it online. "Media convergence" has become a "community" of interests.

2.2 Characteristics of advertising communication in the context of media convergence

Advertising communication in the context of media convergence has the following characteristics:

2.2.1 Resource accommodation.

It is to reasonably integrate the human and material resources of the new and old media,and change their respective services into common services. First of all,the radio and website were merged,the original editors of both sides were connected,and the "media convergence editing center" was established. When the reporters of the center go out to interview,they carry the recording pen and digital camera at the same time to provide contributions for the broadcast and the network at the same time,which not only ensures the source of both sides' press releases,reduces the labor cost,but also improves the authority and original ability of the website's press releases.

2.2.2 Publicity is integrated.

To establish a new type of harmonious,complementary and mutual trust media relationship. To create a "fusion media" is to straighten out the relationship between the new and old media,analyze the advantages and disadvantages of the new and old media,complement each other with advantages,promote the advantages and eliminate the disadvantages,and achieve the effect of $1 + 1 > 2$. For example,the rapidity and convenience of broadcasting,the intuitiveness and stereoscopic of television,and the "four infinities" (infinite space,infinite time,infinite author,and infinite audience) of Internet. The media are not jealous or afraid of the things that "I have nothing else",and they are not exclusive or refuse to the things that "he has nothing else". They all regard "he" as a part of themselves,and pursue the "bucket law".

2.2.3 The interests are in harmony.

The ultimate goal of developing "Media convergence" should be beneficial to the fundamental benefit. The benefits are mainly reflected in two aspects,namely social benefits and economic benefits.

2.3 Function architecture of online directional advertising push service

Based on the TV station resources and many Internet media resources,this paper adopts the strategy of media integration to integrate the single means of communication. In this way,traditional media and new media are integrated. The online and offline media are integrated to produce synergy effect and realize strategic communication mode. At the same time,it can give advertisers the best suggestions to help advertisers achieve the optimal delivery. Accurate we media delivery platform,accurate big data capture and manual review give advertisers the most accurate delivery suggestions as well as reducing delivery detours and maximize delivery effect. This provides customers with high quality and efficient service.

The rapid development of the internet has spawned internet advertising,and the advent of the era of big data has forced advertisers to develop in the direction of Internet targeted advertising. The emergence of internet targeted advertising has been the inevitable result of social development. The purpose of directional advertising is to deliver ads to target user groups,so the network directional advertising push service is born. Network directional advertising push service is a process of matching advertising and user behavior characteristics,which involves four aspects in the whole process of the system: ad-publishers,advertisers,advertising agency and users. Among them,the advertisement publisher registers the directional advertisement service with the advertisement agency,the advertiser puts forward the advertisement demand to the advertisement agency,the advertisement agency comprehensively considers the user interest demand,the advertiser bid and other factors,and finally completes the directional advertisement service to the user on the

advertisement publisher platform(Shen Xiaomin,2001;Cai Qi & Huang Yaoying,2011;Wei Lu & Ding Fangzhou,2013). The four are interdependent,interdependent,and win-win through targeted advertising.

The task of directional advertising push service is to push ads that meet the user's interest. Therefore,according to the overall function of directional advertising push service,the whole service system can be divided into two parts,one is Data Management Platform (D VIII),the other is demand side platform (DSP). As shown in Fig. 1.

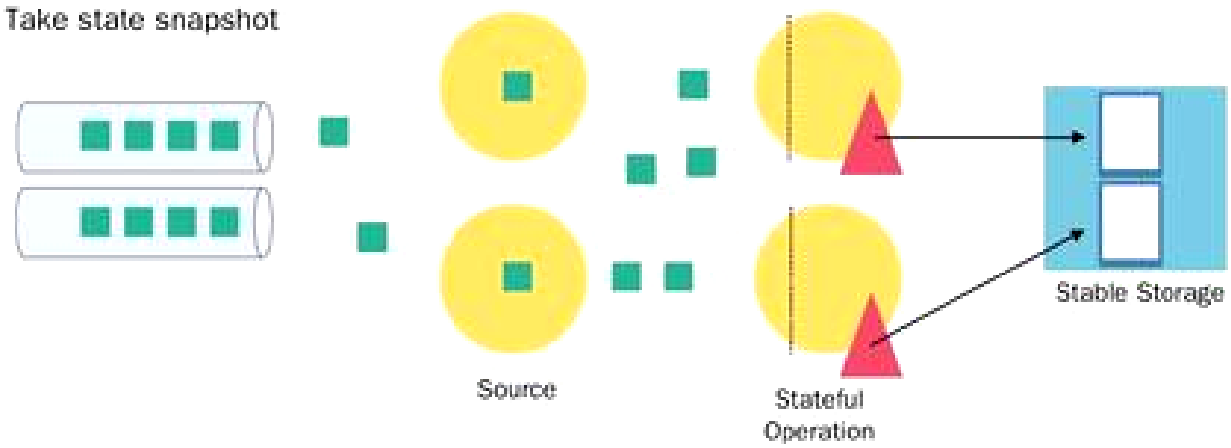

Fig. 1- Function architecture of online directional advertising push service.

Data Management Platform,also known as the central Data Management Platform,the main function is to collect users' online data to this technology platform,then analyze and extract users' online behavior interest,and establish accurate user segmentation. The demand side platform is the direct embodiment of the network directional advertising push service,and its function is to complete the task of putting the directional advertising in line with the user's interest. The two cooperate to complete the network directional advertising push service.

The specific push service process of network directional advertising starts from the acquisition of users' online data,in which the Data Management Platform is responsible for completing this task,and then the Data Management Platform analyzes users' online behavior data through certain methods,such as independent user identification,web keyword extraction,similar web aggregation,behavior analysis,etc. Finally,the user's online interest feature file is extracted. After that,the demand side platform began to play a role,which took the user's online interest profile output by the Data Management Platform as the data input. Then the advertiser's demand is compared with the user's online interest characteristics. Through a certain pricing method and bidding mode,the highest bidder is selected from the ads that meet the user's interest characteristics,and the online behavior of the end user triggers the demander platform to put the ad to the user. The specific process of targeted advertising is shown in Fig. 2.

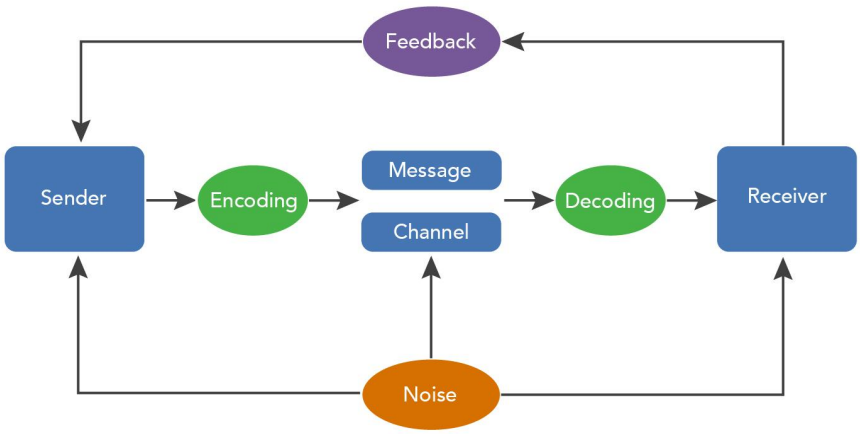

Fig. 2 - Specific process of network targeted advertising.

2.4 Data Management Platform

Data Management Platform (DMP), as a big data platform for users to collect and analyze online data,is the core part of network oriented advertising push service. Without the existence of DMP,it is impossible to extract the online behavior characteristics of users,and then there is no match between advertising and user behavior characteristics,directional advertising can only become empty words. From this,we can see the important position of Data Management Platform in the whole network directional advertising push service.

DMP is also a core component of DSP,which makes the data of advertising audience reasonably controlled by advertisers and advertising agents. Through these data,we can have a more intelligent advertising plan management and media purchase. DMP collects the interest attribute data of Internet users,analyzes,subdivides and stores these data,and finally classifies the user groups in the data according to different interest characteristics. According to the latest survey by IAB and Winter-berry group,more than 80% of advertisers,marketers,publishers,technical developers and marketing service providers believe that DMP will play a key role in improving marketing and advertising. As a comprehensive platform for big data collection,integration and management,DMP will benefit from it,especially from those organizations seeking to improve the effectiveness of advertising objectives.

3. Key technologies of user behavior characteristics analysis

3.1 Hadoop platform

Hadoop is a software framework that can process a large amount of data in a distributed way. In fact,it is built on the cluster composed of many cheap hardware platforms,and it is oriented to the storage of large files. It is a distributed computing platform that allows users to easily develop and run applications dealing with massive data on Hadoop,and it is also a platform that users can easily manipulate and master. It has the advantages of high reliability,high scalability,high efficiency and low cost.

Hadoop framework is composed of many elements,such as Map-reduce,HDFS,HBase,Pig,Zookeeper and Hive,among which the core design is distributed file system HDFS and parallel programming framework Map-reduce. HDFS provides storage for massive data,while Map-reduce provides calculation for massive data. HDFS is at the bottom of the whole Hadoop framework,and it is the foundation of all modules in the framework. HDFS and Map-reduce are two core components of the platform.Next,this paper will introduce the Map-reduce parallel computing framework in detail.

3.2 Map-reduce parallel computing framework

Map-reduce is a distributed computing framework for parallel computing of massive data sets. Its characteristic is that it can reduce the difficulty of parallel computing programming,so that programmers can easily write efficient parallel programs. Map-reduce enables users to pay attention to the processing logic of Map and Reduce related to specific applications when developing parallel computing,and hand over their complex parallel transactions to the underlying system.

The main idea of Map-reduce mode is to solve the problem of automatic segmentation by means of Map and Reduce. First,the program of Map function maps the segmented data into multiple data blocks and distributes them to computer clusters for distributed operation. Then,the program of Reduce function summarizes the operation results,so as to output the results required by developers. The specific processing flow is shown in Fig. 3.

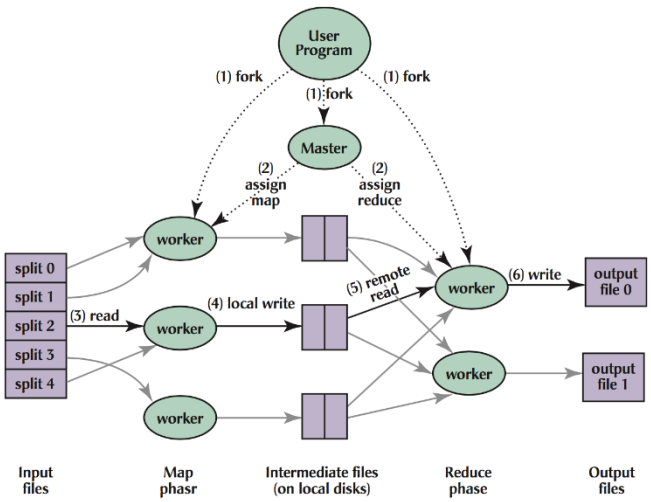

Fig. 3 - The processing flow of Map Reduce.

In the Map stage of the processing process,each Map function receives several independent data blocks,generates a set of key-value pairs < key,value > of intermediate results,and sorts them according to key. In the Reduce stage,all the intermediate data are traversed,and the user-defined Reduce function is executed for each unique key value,and the key-value pair < key,value > is obtained as the output result.

4. Implementation of sub-module of Data Management Platform

4.1 Outline design

The Data Management Platform sub module is an experimental platform built according to the needs of this paper. The platform realizes some functions of the Data Management Platform,including user data preprocessing and interest feature analysis. The user data preprocessing part mainly deals with the user's online log,including key field extraction and independent user identification. User log file is a massive data set,so this part is the process of processing a large number of user data,which is based on Hadoop platform to complete the processing task. User behavior characteristic analysis is the process of analyzing independent users. Its data source is the standardized data after preprocessing. The main purpose is to extract the user's online interest characteristics. The function flow of Data Management Platform sub module is shown in Fig. 4.

Fig. 4 - Data Management Platform sub module function architecture.

4.2 Function module implementation

The main task of data preprocessing module is to extract the key fields of user ID (userld),access URL (URL) and web page time (time) from the semi-structured log file,and to identify the independent Internet users. The data source of this system is the network log of the information center of Beijing Jiaotong University. Each student only uses one account to access the Internet in the school. Therefore,according to the actual situation of the data source and the needs of this paper,it only depends on the user's online account to identify independent users.

Firstly,a Map function is used to process the input log file,and each piece of data is Record;UserlD,Time stamp of web page and URL of visit are intercepted,and the output result of Map process is Keyi (Valuei). In which

Keyi=UserlDi,Valuei=Listi (<Timei,URLi >),and then the output result of the Map process is taken as the input of the Reduce process. the purpose of the Reduce function is mainly to process the Value of the input key-value pair. And each Value is a List array,and the elements in the array are the combination of Time and URL. the task of Reduce is to sort the Values according to the order of time from the new to the old. in the output results,Keyi=UserlDi,valuei*= Listi* (< Timei,URLi >),and finally each different Key outputs the results according to the time-ordered rules in value.

The task of the web page feature word extraction module is to extract the words representing each web page feature. After data preprocessing,each online log is sorted into records including URL,Time and UserlD. The extraction of web page feature words mainly deals with URL fields,and the specific process includes web page content grabbing,Chinese word segmentation,stop words filtering and TF. IDF calculating vocabulary weight values.

4.2.1 Web page content capture

Web content crawling is realized by jsoup Web crawling tool. Considering the structure and content distribution of each web page,only the Title,Keywords and Description are captured,which basically cover the key topics of the web page. If you continue to capture other contents of the web page,it will interfere with the extraction of web page features. Considering the importance of these three parts,they are weighted respectively,in which Title is 4,Keywords is 2,and Description is 1. If Keywords and Description exist in the web page,it will be output directly. If they do not exist,they are replaced by "Null Description" and "Null Keywords" respectively,so as to effectively identify them in the subsequent process.

4.2.2 Chinese word segmentation and stop words filtering

The main task of this module is to segment the captured web page content in Chinese,that is,to cut the text content in Title,Keywords and Description into single words. Chinese word segmentation adopts the ICTCLAS open source system,which is a Chinese lexical analysis system based on multi-layer hidden codes developed by the Institute of Computing Technology,Chinese Academy of Sciences. This paper adopts the open source jar package of the system,and the word segmentation effect is good after testing. In stop words filtering,a stop-list containing 4546 stop words is adopted,which effectively removes useless words such as "moreover","de" and "if"

5. Experimental results and analysis

5.1 Experimental data selection

The experimental data required in this paper mainly includes two parts,namely user log and text advertisement. The users' online logs are from the network logs of information center of Beijing Jiaotong University,and the 10 day network logs from April 15 to April 24 are selected as the experimental data. In these massive network logs,there are 400 online records generated by "13125154" in ten days. Among the 400 online records,there are many web pages on the topics of tourism,rental,stock,automobile,NBA,English training,skin allergy,etc. among them,there are 35 tourism web pages,mainly distributed on the 23rd and 24th. There are 30 rental web pages,which are concentrated on the 21st,22nd and 23rd; 40 stock web pages,which are evenly distributed in 10 days; 35 NBA web pages,which are evenly distributed in 10 days; and 50 English training web pages,which are mainly distributed in 17,18,19 and 20 days. The results showed that there were 30 pages for automobile,25 pages for skin allergy,25 pages for noise,and 30 pages for automobile,25 pages for skin allergy,15 days,16 days and 17 days. All the text ads come from the sponsor ads of real commercial search engines,including 330 theme related ads and 150 theme independent ads as noise.

5.2 Analysis of experimental results

5.2.1 Analysis of similar web page aggregation optimization effect

In this experiment,firstly,the online log in the test data is preprocessed,and three key fields of UserlD,Time and URL in the log are extracted,and the user whose online account number is "13125154" is identified,and then the behavior characteristics of the user are analyzed independently. The size of the tracking window 121 is set to 400,and the behavior standard is set to 3%. The similar web page aggregation methods before and after optimization are applied to the analysis of user behavior characteristics,and then the web pages in the tracking window are analyzed respectively. Then,from the two kinds of analysis results,we select the web page clusters with the characteristics of travel,renting,stock,automobile,NBA,English training and skin allergy. Finally,we compare and analyze the accuracy and recall rate of similar web page aggregation,and finally verify the optimization effect of similar web page aggregation.

5.2.2 Analysis on the improvement effect of advertising matching accuracy

Through the analysis of users' online records by Data Management Platform,multiple interest features of users are finally obtained. Each interest feature is composed of a web page set,and each interest feature has a corresponding weight value. If only feature freshness is considered and feature dispersion is not considered,then the discretization factor is set to 0. On the 24th,all 400 web pages as experimental data entered the tracking window. At this time,the weight value of each feature was the value weight of each user's interest feature on the 24th. Tourism web pages were just visited on the 23rd

and 24th. Therefore,the value weight of this feature is the highest,and the value weight of skin allergy web page and automobile web page is relatively low because they visit earlier. Because the feature dispersion is not considered,the weight value of stock web pages and NBA web pages distributed evenly in the tracking window is not particularly high. According to the analysis results,we can find the ads that match the interest features with high value weight in the advertising database,and then put them to users,in order to increase the click through rate of related ads. The process of matching user interest features and advertisements is the same as the principle of similar web page aggregation. Integrating the Redis based web page category determination method into the process of advertisement matching can also optimize the effect of matching advertisement and user interest features.

6. Conclusion

This paper mainly verifies the application effect of Redis based web page category determination method in online advertising push service. Firstly,the experimental data is selected and the experimental evaluation index is set,and then the optimization effect of similar web page aggregation optimization strategy is verified and demonstrated. Then,referring to the similar web page aggregation optimization strategy,the Redis based web page category determination is integrated into the process of user interest feature and advertisement matching to form an advertisement matching optimization method. Finally,the improvement effect of the optimization method on the accuracy of user interest and advertisement matching is verified.

Acknowledgement

The authors would like to thank the fellow authors and organizations whose intellectual properties were utilized for this study.

Conflict of Interest

The authors declare no conflicts of interest.

References

Cai Qi, & Huang Yaoying. (2011). New Media Communication and the Development of Audience Participation Culture. Journalist, (08):28-33.

Chen Hong, & Liang Junmin. (2013). Opportunities and Challenges of Health Communication Development Under the New Media Environment. Journalist, (05): 60-65.

Deng Xinmin. (2006). We Media: the Latest Stage and Characteristics of New Media Development. Exploration, (02):134-138.

DJiang Enlai. (2009). Ideological and Political Education of College Students Under the New Media Environment. Reprinted Newspaper Materials: Ideological and Political Education, (06): 59-61.

Liu Haiye. (2022). Research on the Transformation of Journals of Local Ethnic Colleges under the Background of Media Convergence. Academic Journal of Humanities & Social Sciences, 5 (18):40-44.

Ma Deyong, & Sun Mengxin. (2014). The Determinants of Government Credibility in the New Media Era: Transparency, Responsiveness or Public Relations Skills. Journal of Public Management, 011(001): 104-113.

Shen Xiaomin. (2001). Research on the Nature and Function of Textbooks in the New Media Era. Global Education Outlook, (03):23-27.

Sun Shaoyi. (2008). Cultural Transformation and the Role of Intellectuals. Journal of Shanghai University: Social Sciences, (03):33-38.

Wang kun kun. (2014). Media and Collective Memory Research: Review and Reflection. Journalism and Communication Research, (09): 39-50.

Wei Lu, & Ding Fangzhou. (2013). On the Transformation of Communication Research in the New Media Era. Journal of Zhejiang University (humanities and Social Sciences), 43(4): 93-103.

Xu Xiang. (2022). Construct a new pattern of digital education management in colleges and universities in the era of media integration. International Journal of New Developments in Education, 4(09):91-95. https://dx.doi.org/10.25236/IJNDE.2022.040915

Yang Xiaohui. (2009). New Media and New Ideas of Ideological and Political Education. Theoretical Front of Colleges and Universities, (07):43-44.