Digital Reading Recommendation Model Based on Fast Data Processing Technology

Xiuxian Li[1](✉) and Rasha Almajed[2]

[1] Tianhua College, Shanghai Normal University, Shanghai 201815, China
happylixiuxian@126.com
[2] American University in the Emirates, Dubai, UAE
rasha.almajed@aue.ae

Abstract. Fast data processing is one of the important technologies and development directions of big data processing, which fully embodies the characteristics of big and fast big data. The digital reading recommendation model based on fast data processing technology can make full use of the advantages of big data, and further improve the digital reading experience of readers.

Keywords: Fast data processing · Digital reading recommendation · Big data environment

1 Introduction

In the history of human civilization, reading plays an important role. With the development of Internet information technology and the emergence of new digital carriers, people's reading methods are changing. Digital reading has gradually become a new trend of people's reading in modern information society. Digital reading includes the digitization of reading objects, such as the transformation from traditional paper and ink to e-books, web pages, blogs, etc.; It also includes the digitization of reading mode. The carrier of reading is transformed from flat paper into electronic digital devices such as computers, mobile phones and mobile terminals.

In the fast-paced modern society, information often has timeliness, such as news, microblog, instant messaging and so on. If we can deal with these real-time information quickly and provide real-time data response services, we can effectively meet the needs of users' real-time information. Cloud service is built on the MapReduce framework of parallel processing, which often needs a large time delay and can not respond to the data processing tasks with high real-time performance in time, which leads to the lack of timeliness of data processing and reduces the utilization value of real-time data. This paper consists of the following parts. The first part introduces the related background and significance of this paper, the second part is the related work of this paper, and the third part is data analysis. The fourth part is example analysis. The fifth part is conclusion.

I. Ahmad et al. (Eds.): STSIoT 2021, LNDECT 122, pp. 679–687, 2023.
https://doi.org/10.1007/978-981-19-3632-6_79

2 Related Work

Kucirkova reviewed the personalization logic embedded in reading recommendation systems developed for 2- to 11-year-old children and its (dis)alignment with Papert's constructionist and socio-constructionist theories of learning [1]. Ref [2] conducted an empirical analysis to examine the interaction effects of these three decision-supporting tools on online reading behavior. Orero et al. described some further standardisation requirements such as terminology, intended audience, workflows, formats, and languages that should be taken into consideration towards a 21st century Easy to Read recommendation [3]. In order to recommend literary works of interest to readers and improve readers' reading efficiency, an intelligent recommendation method for literary reading based on user social network analysis is proposed [4]. Lv et al. proposed a probabilistic generative model, BoRe, where user interests and crowd effects are used to adapt to the instability of reader consumption behaviors, and reading sequences are utilized to adapt user interests evolving over time [5]. In order to provide accurate and effective book sets for users Yu et al. proposed an algorithm based on fusing their preferences [6]. Readability analysis of 405 identified HEMs revealed scores above the sixth grade reading level recommendation [7]. To narrow these research gaps, Zhang et al. conduct a diary study to capture a comprehensive picture of readers' use of algorithm- and social-sourced information to inform their future reading choices [8]. Ref [9] use dynamic density clustering method and stream computing based on time series analysis to give the label system a time dimension by combining the big data characteristics of users. Other influential work includes Ref [10].

Based on the above research status at home and abroad, the author proposes a digital reading recommendation model based on fast data processing technology to provide users with real-time and efficient digital reading recommendation service, stimulate users' interest in reading, and meet the requirements of users in the rapidly changing modern society.

2.1 Types of Digital Reading

With the advent of 3G era, 2009 is known as the "e-book year" in the world. Foreign Amazon's Kindle, Sony Reader of Sony and iRiver story of Sony, and domestic Hanwang, Shanda, Hanlin, Patriot oppo and other readers are widely concerned. Digital reading such as computer reading, mobile reading, e-book reading and so on has become popular. The type flow of digital reading is shown in Fig. 1 below.

(1) Computer reading. Computer reading refers to a kind of reading behavior that uses multimedia computer technology and network technology to obtain information and knowledge and complete meaning construction. Compared with the traditional reading, the reading text of the main body has no qualitative change, what it changes is the presentation and use of the text. According to people's reading needs, there are mainly two types of computer reading: one is real-time online reading, which refers to reading the network resources collected by readers directly on the Internet and always hanging on the Internet in the process of reading. The second is to download offline reading, which means that readers download network resources to another

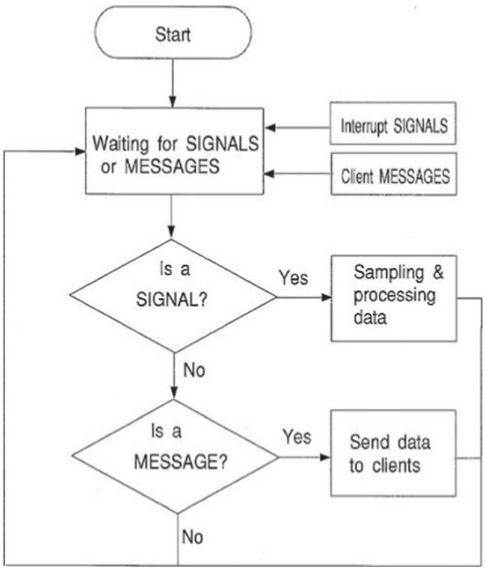

Fig. 1. Type and process of digital reading.

carrier and then read after breaking the network link. There is also a hybrid reading mode combining network reader and traditional reading mode, that is, downloading and printing network content for reading. There are two main carriers of computer reading. One is PC or terminal; The other is a special reader, e-book.

(2) Mobile reading. Mobile phone reading refers to the reading activities which use mobile phone as the terminal, access, accept and download the required information through mobile communication network, and browse, watch (listen) on the mobile phone. As a reading carrier, mobile phones need to have the following functions: interactive information receiving terminal; Self communication terminal; Information carrying terminal; Identity confirmation terminal and data acquisition terminal. Mobile reading enables users to make use of some fragmented time, which makes up for the coverage of traditional reading methods, reduces the threshold of reading to a certain extent, increases the reading user group, and is a popular reading method, as shown in Fig. 2.

(3) E-book reading. E-book is the abbreviation of e-reader. It is a new digital reader with e-paper as the display screen. It can read most e-books on the Internet, such as pdf.chm.txt and so on. It is coated with an electronic ink composed of numerous tiny transparent particles. As long as there is an electric field, and the electric field action mode changes, the display image can be changed. Although the physical basis of information representation is different, the electronic ink display has no obvious difference with paper in use, and can even be folded and rolled up like a newspaper. Digital code folding calculation formula:

$$\|\Delta x_{k+1}(t)\| \leq \int_0^t e^{(pk_f+m_2+m_3)(t-\tau)}(m_1\|\Delta u_k(\tau)\| + pd)d\tau \tag{1}$$

Fig. 2. Mobile digital reading

Inequality:

$$\|\Delta x_{k+1}(t)\| \leq \left(pk_f + m_2 + m_3\right) \int_0^t \Delta x_{k+1}(\tau) d\tau + \int_0^t (m_1 \|\Delta u_k(\tau)\| + pd) d\tau \quad (2)$$

2.2 Advantages of Digital Reading

(1) It has a large amount of information and rich content. Rich reading content mainly comes from the network itself and its links. Because the amount of information that the storage space of the computer can carry is unmatched by the paper carrier, and at the same time, it can exchange information with the host computers of different countries and regions across time and space restrictions, which makes the network become a huge information resource database. Network content updates quickly, can grasp the spirit of the times in time, can reflect the latest trend of social development and change in time, make people feel the breath of the times.

(2) Strong interaction and open environment. Hypertext reading makes it possible for people to communicate with computers. Retrieval is convenient. Browsing and jumping reading make readers have a broad mind, a strong desire to participate, and reading becomes more fun.

(3) The display is lifelike and easy to carry. The advantages of digital reading media in convenience and large capacity are incomparable to paper reading media. Digital reading media takes up less physical space and has a large storage capacity (Ig capacity can store 536870912 words); E-book is similar to paper comfort, lifelike display effect, to meet the requirements of people's comfortable reading anytime and anywhere_ It's a mobile library on the Internet.

(4) More environmentally friendly, more economical. At present, the price of e-books on the market is mostly between 1000 yuan and 2500 yuan. Although the one-time cost is much higher than that of paper books, the storage capacity of e-readers is quite large and the content can be updated. There is no need to publish new books

like paper books, so it is necessary to re paper, typeset and print, which consumes resources. In the long run, the cost of e-books is far lower than that of paper books, It can improve and protect the environment more effectively.

(5) Digital reading makes ubiquitous reading a reality "Ubiquitous reading" is ubiquitous instant reading, that is, people can read the latest text anywhere. With the support of all kinds of new information technology, people can easily read e-books on the beach, forest and desert. In March 2010, a French (Figaro) survey showed that French paper reading is still the mainstream, but 22% of the people are ready to accept digital books. Hanwang Technology's domestic test in 2010 showed that only 5% of Internet users wanted to experience e-book at the beginning of 2009, while 65% wanted to have an e-book terminal by the end of 2009. With the popularity of e-books, these data seem to be suggesting that the reality of ubiquitous reading is getting closer and closer to us. As shown in Fig. 3, the number reading relationship.

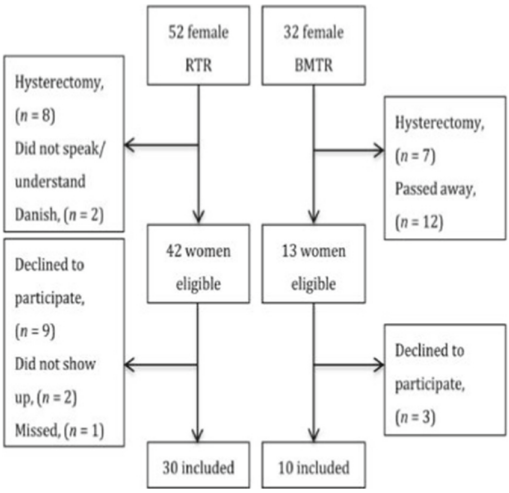

Fig. 3. Digital reading relation

(6) Digital reading can be light reading or deep reading. We can browse on the Internet, or read the excellent literature carefully. At the same time, what is more convenient than traditional deep reading is that we can also mobilize all kinds of online resources, including noun explanation, book review, video and audio materials, to assist deep reading.

(7) Digital reading supports readers' reading habits, protects readers' Reading Rights, and saves time and cost. In digital reading, libraries, digital publishers and sharing websites jointly build a "cloud" of digital reading. No matter what you want to read, you can go to the "cloud" to get it, which not only ensures the reading rights of readers, but also saves time and cost.

2.3 Real Time Data in Digital Reading

At present, the number of Internet users in China has exceeded 564 million 4. Instant network services such as personal blog, microblog, social networking website and instant news are widely accepted by everyone. The network era of instant application has come. When users use real-time digital reading applications, such as mobile library, blog, microblog, wechat and real-time news push tools, they will produce a large amount of real-time data, which reflects the changing trend of user needs in real time and dynamically. In the digital reading service, by collecting, sorting, mining and analyzing these real-time data, we can timely understand and analyze users' reading psychology and habits, and timely adjust the digital reading recommendation service strategy according to changes, so as to quickly meet the needs of readers.

The real-time data in digital reading has the characteristics of real-time, volatile, sudden and infinite, which reflects the "fast" data characteristics of big data. These real-time data have significant timeliness and need to be processed immediately, otherwise most or even all of the data application value will be lost. Therefore, the real-time data acquisition in digital reading is often completed through a specific time window or predetermined data acquisition conditions, which can not only meet the needs of rapid real-time data acquisition, but also obtain valuable application data in a large number of dynamic data streams. In the analysis of real-time data, efficiency often plays a more decisive role than effect. Real time data applications often need high-speed data processing and response speed, can quickly respond to data processing tasks in seconds, and provide users with more accurate data processing results.

3 Data Analysis

3.1 Delay Data in Digital Reading

The construction of digital reading in China has gone through more than ten years. In the process of digital construction, libraries and digital publishing institutions have various types of academic resource databases, digital e-books, electronic journals, library databases, which contain a lot of bibliographic information. In addition, there are various metadata and ontology describing these bibliographic information, such as RDA data format, FOAF ontology, etc., which are gradually accumulated delay big data. They are the most important part of digital reading. These resources are of various types, huge amount and in the process of continuous growth.

These kinds of delay data in digital reading often have the characteristics of massive, high value, data generation needs to be accumulated, which is in line with the "big" of big data, that is, the characteristics of large amount of data, great value and great demand. It is an indispensable part of digital reading application, and also the data basis of digital reading recommendation service. For these delayed data in digital reading, it is often necessary to carry out batch big data processing and continuous data collection, not only to save historical data, but also to add new data information regularly. Therefore, the delay data in digital reading need reliable and stable data acquisition and storage, in order to ensure that the data can be fully and effectively used.

$$\Delta x_{k+1}(t) = \int_0^t P^{-1}(\tau)(f(t, x_d(\tau)) - f(t, x_{k+1}(\tau)))d\tau \tag{3}$$

$$\Delta x_{k+1}(i) = \int_0^t P^{-1}(\tau)\Delta x_{k+1}(\tau)d\tau - \int_0^t P^{-1}(\tau)d_{k+1}(\tau) \quad (4)$$

3.2 Real Time Data in Digital Reading

At present, the number of netizens in China has exceeded 564 million. Real time network services such as personal blogs, microblogs, social networking sites and instant news are widely accepted. The era of real-time applications has come. When users use real-time digital reading applications, such as mobile library, blog, microblog, wechat and instant news push tools, they will produce a large amount of real-time data, which reflects the changing trend of users' needs in real time and dynamically. In the digital reading service, by collecting, mining and analyzing these real-time data, we can timely understand and analyze the reading psychology and habits of users, and timely adjust the digital reading recommendation service strategy according to the changes, so as to quickly meet the needs of readers. The real-time data in digital reading has the characteristics of real-time, volatile, sudden and unlimited, which reflects the "fast" data characteristics in big data. These real-time data have a significant timeliness, need to be processed immediately, otherwise it will lose most or even all of the data application value. Therefore, the acquisition of real-time data in digital reading is often completed through a specific time window or predetermined data acquisition conditions, which can not only meet the needs of rapid acquisition of real-time data, but also obtain valuable application data in a large number of dynamic data streams.

4 Example Analysis

4.1 Overview of Fast Data Processing Technology

Big data contains data with high production speed, such as click stream data, financial transaction data, log aggregation data or sensor data. These events often occur thousands or even tens of thousands of times per second, reaching MB per second, GB per hour or TB per day. The best way to capture the value of input data is to react and operate immediately when the information arrives. If the input data is processed in batch, it needs a lot of delay, which means that it may lose its timeliness, and then lose the core value of fast data. In order to solve the problem of real-time data processing in big data environment, researchers put forward stream computing technology, which can efficiently process the fast generated data stream in big data environment. Stream computing technology regards the fast generated real-time data as the form of data stream, does not store and calculate all the data, and does not need to determine the time and order of data generation. Instead, it directly performs real-time operation on the data that meets certain windows or constraints in the data stream, and outputs the results quickly. Although stream computing technology can solve the problem of real-time data processing in big data environment, because stream computing does not store historical data, data processing has the characteristics of disorder and volatility, so the accuracy of data processing can not be comparable with batch data processing. At the

same time, due to the lack of data context environment, some complex data processing algorithms and technologies such as natural language processing and data mining can not run effectively, which can not meet the needs of complex user data processing.

Fast data processing technology is a new big data processing framework developed after cloud computing technology and stream computing technology, which integrates cloud computing technology with stable and accurate delay processing of big data and stream computing technology with fast and reliable real-time processing of real-time data. It has been adopted by mainstream network application companies such as Google, and gradually replaces the traditional cloud computing processing platform, It has become a new trend of big data processing technology.

4.2 Digital Reading Recommendation Framework

The author proposes a digital reading recommendation framework based on fast data processing to meet the needs of users for fast and real-time digital reading recommendation, and provide users with accurate, efficient and comprehensive digital reading recommendation services.

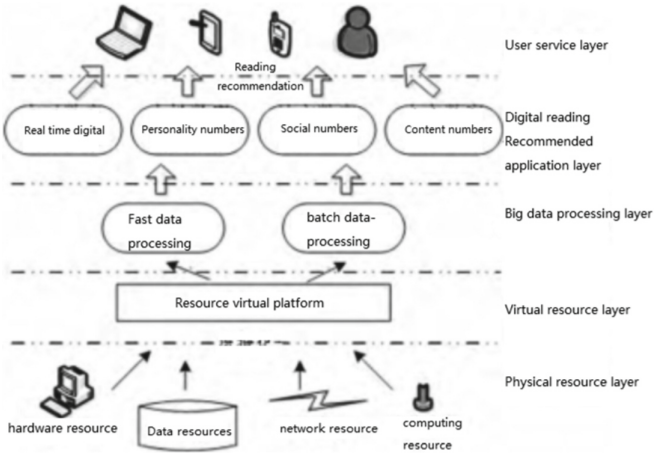

Fig. 4. Digital reading recommendation framework based on fast data processing technology

As shown in Fig. 4, the digital reading recommendation framework is mainly divided into five parts: physical resource layer, virtual resource layer, big data processing layer, digital reading recommendation application layer and user service layer. Among them, the physical resource layer and virtual resource layer are the big data processing layer; Data processing layer and digital reading recommendation application layer are the core parts of the framework and the innovation of this paper, which provide users with digital reading recommendation service based on fast data processing technology; The user service layer can subdivide different users to meet the needs of mobile users, traditional PC users, convenient device users and other users for digital reading recommendation. These five levels are progressive and complement each other to provide users with comprehensive, efficient and fast digital reading services.

5 Conclusion

This paper discusses and analyzes the digital reading recommendation framework based on fast data processing technology, and puts forward the digital reading recommendation framework based on fast data processing technology, which lays the technical foundation for the development of digital reading recommendation service in the big data environment, and puts forward the preliminary construction scheme. In the future research and practice, the author also needs to further study its specific implementation and application, learn from the existing technical experience of fast data processing application service, integrate the existing digital reading service in the big data environment, optimize and improve the digital reading recommendation service, in order to further improve the efficiency and effect of digital reading recommendation service, To meet the needs of users for efficient, accurate and comprehensive digital reading recommendation service in the big data environment.

References

1. Kucirkova, N.: The learning value of personalization in children's reading recommendation systems: what can we learn from constructionism? Int. J. Mob. Blended Learn. **11**, 80–95 (2019)
2. Liu, Q., Zhang, X., Zhang, L., Zhao, Y.: The interaction effects of information cascades, word of mouth and recommendation systems on online reading behavior: an empirical investigation. Electron. Commer. Res. **19**, 521–547 (2019)
3. Orero, P., Delgado, C., Matamala, A.: Easy to read standarisation. Some steps towards an international standard. In: 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion (2020)
4. Han, Q.: Intelligent recommendation method of literature reading based on user social network analysis. In: 2020 International Conference on Robots and Intelligent System (ICRIS) (2020)
5. Lv, P., Meng, X., Zhang, Y.: BoRe: adapting to reader consumption behavior instability for news recommendation. ACM Trans. Inf. Syst. **38**, 1–33 (2020)
6. Yu, Z., Bu, J., Li, S., Wang, W., Tang, L., Zhao, C.: Research on book recommendation system for people with visual impairment based on fusion of preference and user attention. In: Miesenberger, K., Manduchi, R., Covarrubias Rodriguez, M., Peňáz, P. (eds.) Computers Helping People with Special Needs, vol. 12376, pp. 83–90. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58796-3_11
7. Zhang, H., Buchanan, G., McKay, D.: Hey Alexa, What should i read? Comparing the use of social and algorithmic recommendations for different reading genres. In: Toeppe, K., Yan, H., Chu, S.K.W. (eds.) Diversity, Divergence, Dialogue, vol. 12645, pp. 346–363. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-71292-1_27
8. Correa, D.J., et al.: Quantitative readability analysis of websites providing information on traumatic brain injury and epilepsy: a need for clear communication. Epilepsia **61**, 528–538 (2020)
9. Liao, J.: Design of library user profile system based on dynamic density clustering algorithm and stream computing. In: 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) (2021)
10. Sun, H.: Research on interest reading recommendation method of intelligent library based on big data technology. In: Web Intelligence (2020)