Key Technology Review of Big Data Analysis in the Social Network

Zhendong Ji[1]

[1] Shandong Women's University, 250300 Jinan, China
Namejzd@ 163.com

Abstract. As the network promoting in people's lives, more and more people use the social networks and other network platform, the concepts of big data have been gradually reference to the network data analysis, researching big data analysis framework and technology can improve enterprise management level. This paper analyzes the framework and its key technology of big data analysis in the context of social network, in order to be able to provide the corresponding theoretical support and reference for analysis of social network data.

Keywords: Big data analysis, social networks, key technology

1   Introduction

With the continuous development of the Internet, a variety of social networks emerging, there are more and more data on the social network and gather into big, analyzing in depth of these data can effectively resolve various issues of social and enterprises, [1] and big data analysis provides a convenient for enterprise to analyze network data, there are more and more institutions adding into the research of big data analysis, the development of big data analysis techniques has brought enormous value for social management and enterprise management.

Because China's technology of information network falls behind western developed countries and starts late, related personnel need to do corresponding theoretical study and specific model analysis and study of the key technology based on the existing characteristics of big data in China according to the analytical model of big data in advanced countries, in order to promote the analysis of large data formation in China in the framework of the society under the background of the network.

2   The Framework of big data analysis

In the context of social networks, big data analysis technology has become a hot technology in network data analysis, social network data analysis covers anthropology, sociology, psychology, geography, communications technology,

organizational study, biology and other aspects of knowledge, it is a multi-disciplinary cross multifaceted technology, the social network also contains a large information data, the relationship data and user data, etc. to analyze the different types of data requiring complex analytical techniques to accomplish, the corresponding data analytical framework can structure for all aspects of a comprehensive data analysis[2]. The framework of big data analysis is data layer, the support layer, the analysis layer and the result showing layer, the basic of the whole data analysis is data layer, including capture and storage data for the social network, the analysis layer and the support layer constituted the core of the data analysis framework which analyzing and processing data in-depth, the result showing layer can show data in a clearly and friendly way and apply it to the associated program which, specifically, the functionality and structure of each layer are as follows:
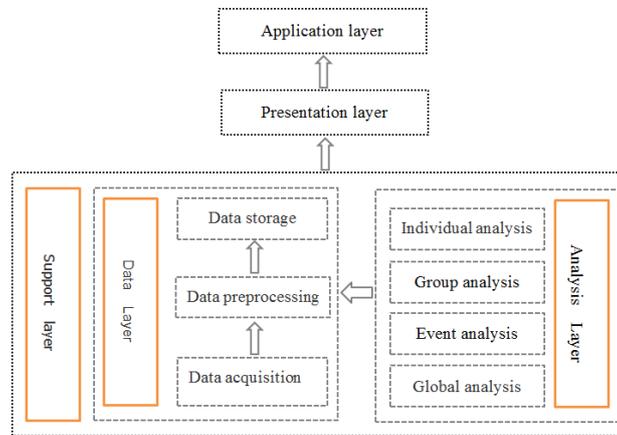

Fig. 1. Large data analysis framework

2.1 The data layer

The data layer is the basis of the data analysis framework, which includes data acquisition, there are three main ways of the data acquisition by framework, the first way is through the network flow, to get the data by identify and resolve network flow, the second way is by API or not-API, including web analytic method, distributed crawler method and API overrun determination methods, the third way is to get the data by other ways, such as meta-search method, incremental method or the acquisition of specific groups, etc. To collect the various data in social networks through these methods, and then pretreated. The final preprocessing step is processing data association, including data storage, log storage and characteristics storage.

2.2 The analysis layer

The analysis layer is one of the core layer of the data analysis framework, which is further analysis for the preprocessed of data layer, specifically to analyze by the four dimensions, the first dimension is the individual analysis, including analyzing the individual circles, emotions, interest, behavior and characteristics, tightness and area networks, etc., the user's personal interests and behavior habits can by able to understand from the analysis of this information. The second dimension is the group analysis, including analysis of specific groups, analyzing opinions of the group leaders, analyzing the relationship, analyzing groups interest, analyzing the evolution of the group, analyzing the potential group members recommendation, it is possible to understand the group by analyzing the information, and which can manage the group of targeted. The third dimension is the event analysis, including analysis of the events discovery, to restore the path of the incident, to hunt for the source of the event, to analyze the propagation of events, to analyze the opinion leader of the events, to predict the trend of events spread etc. through these analysis, The occurrence and the development of the event can be grasped, which have an important role for the master in the developments of the event. The fourth dimension is the overall analysis, including sorting of the popular characters or events, doing statistics of the overall platform, analyzing the regional hot events, analyzing the global topology, The whole social network can be grasped by analyzing these data, in order to understand the status and the role of social networking in the network groups.

2.3 The support layer

The support layer is another important core of data analytical framework, including distributed parallel processing, machine learning, data mining, visualization techniques and natural language processing technology, these key technologies support the operation of data analysis of the entire social network, analyzing and processing by a intelligent and efficient by these techniques for the data in social network, and getting the relevant analytical data.

3 The research of key technologies in big data analysis

The all levels of data analysis framework need to various technology to support to complete the data collection, processing and showing, these key technologies occupies an important position in the data analysis, the key technology in each level is different, these different key technologies support the entire data analysis framework together.

3.1 The data layer

The key techniques or methods used in data layer consists of three parts, the first one is data acquisition, data acquisition methods mentioned above is contained three

methods, network flow is mainly on the way to obtain data which is not encryption. This technique can get the network flow data quickly by resolving and identifying, the system firstly needs to analyze the corresponding network flow during acquisition, and then extracts the significant network flow characteristics, identifying the target flow from the background flow. Then, resolving the target flow, extracting the relationship data, basic data, forwarding data, etc., to realize data acquisition. The second way is to get the API or not API method to obtain data, API is an application programming interface, there are many social networking services provide API, data analysis framework can obtain the relevant data from that, OAuth1.0 and 2.0 are two common ways, the basis information and relationship data of social network can be obtained by connecting the user interface on, there are a certain number and speed restrictions to protect user data in API providing by social network, to avoid being illegal or violent acquisition, when getting data by API it is necessary to control the frequency and speed, in order to ensure the data acquisition is normal. In order to break through the limitations of API, and now there are non-API acquisition mode like the web resolving, using the technology of Web crawler to simulate user logs to obtain the appropriate data.

3.2   The data layer

The key technical in the analysis layer mainly include individual analysis techniques, group analysis, event analysis and the overall analysis, individual analysis techniques is mainly analysis on users' personal information data, including knowledge of individual emotion, behavior, such as analyzing the individual circles is analyzing the relationship, and then summarized the impact of the individual in the circles, and also be able to predict the evolution of the entire circle through individual circle analysis. Sentiment Analysis is to analyze the user's emotional tendencies, including individual negative emotions, neutral emotions and positive emotions, etc., in order to understand the individual's emotional changes, or the attitude of the individual to the events. Group analysis is analyzing the group which is composed of multiple users, including cohesion among this group, the relationship, the relationship between the groups, interests, behavior, psychological and impact, to understand the organization's position and characteristics in the social network.

4   Conclusion

With the development of the Internet, The scale of social networks have been expanding, and contains more valuable information, in order to explore this information more effectively, it is necessary to do some research and explore on big data analysis techniques. The paper focuses on big data analysis framework and the related key technologies to discuss, hoping to play a valuable role in helping mining the potential value in the social network, to be able to build more social network applications humane and fit people's lives needed based on data analysis.

References

1. Dou, B.L., Li, S.S., Zhang, S.Y.: The social networks analysis by the structure. The Computer Journal, vol. 4, (2012)
2. Yi, C.Q., Bao, Y.Y., Xue, Y.B.: Social networking big data analysis framework and its key technologies. ZTE technology, vol.1, (2014)