The Classic Application Cases of Big Data

Lina Yuan [a], Huajun Chen [b]

College of data science, Tongren University, Tongren, 554300, China

[a]893422817@qq.com, [b]390832683@qq.com

Abstract

With the rapid development of Internet technology, especially in recent years, cloud computing, Internet of things, social networking services, etc., newly-developing services to promote data types and scale of the human society is growing at an unprecedented speed, the era of Big data is officially coming. How to better manage and use Big data has become a universal attention topic. This paper analyzes the background and basic features of big data, focuses on expounding the application cases of the Big data, and finally sums up the new challenges faced by the era of Big data.

Keywords

Internet Technology, Cloud Computing, Internet of Things, Big Data.

1. Introduction

The arrival era of "Big data" is put forward by a global well-known consulting firm-McKinsey. The definition of “Big data” is given by the McKinsey global institute that a sort of scale is large enough to beyond the capability of the traditional database software tools of the data acquisition in the aspect of the acquisition, storage, management and analysis, which has four characteristics of volume data scale, fast data transfer, various data types, and low value density. According to McKinsey, data has penetrated into every industry and the business functional areas today, and become an important productive factors, and the mining and application of massive data heralds a new wave of productivity growth and the advent of the wave of the consumer surplus. While the definition of "Big data" given by the research institution-Gartner is that Big data needs the new processing mode, which has the ability of better decision-making, insight-finding, and the process optimization to adapt mass, high growth rate, and diversified information assets. On into 2012, the word-Big data is increasingly being mentioned, which is used to describe and define the information explosion times produced huge amounts of data, and is named to the related technology development and innovation. It has been on the column cover of New York times and the Wall Street journal, into the White House website news, appeared in some domestic Internet theme lectures in the salon, and even into the recommended investment report written by the sharp-nosed Sinolink securities, Guotai junan, galaxy securities, etc [1-5].

Datum are rapidly expanding and being larger, which determines the future development of the enterprise, while many companies may not have realized the data explosive growth cause problems hidden trouble, but as time goes on, people will be more and more aware of the importance of data for the enterprise. As the New York times wrote in a column published in February 2012, the era of "Big data" has arrived, the decisions will be increasingly made based on data and analysis, and not based on experience and intuition, in the commercial, economic and other fields. In 2013, the data stored in the world is expected to reach about 1.2 ZB (about 1.2 billion TB) bytes, if all these data are printed to be books, these books can cover the entire American for 52 times, if they are stored in the standard CD, these CDs can stack into five piles, and each pile can reach to the moon.

A professor of sociology-Gary king at Harvard University has ever said: "this is a revolution, huge data resources makes each field begin the process of quantification, and whether academic, business or government, all areas will start this process". According to search EI database, we can obtain some

interesting datum from perspective of author and author' affiliation about Big data as shown in Fig. 1 and Fig. 2, respectively.

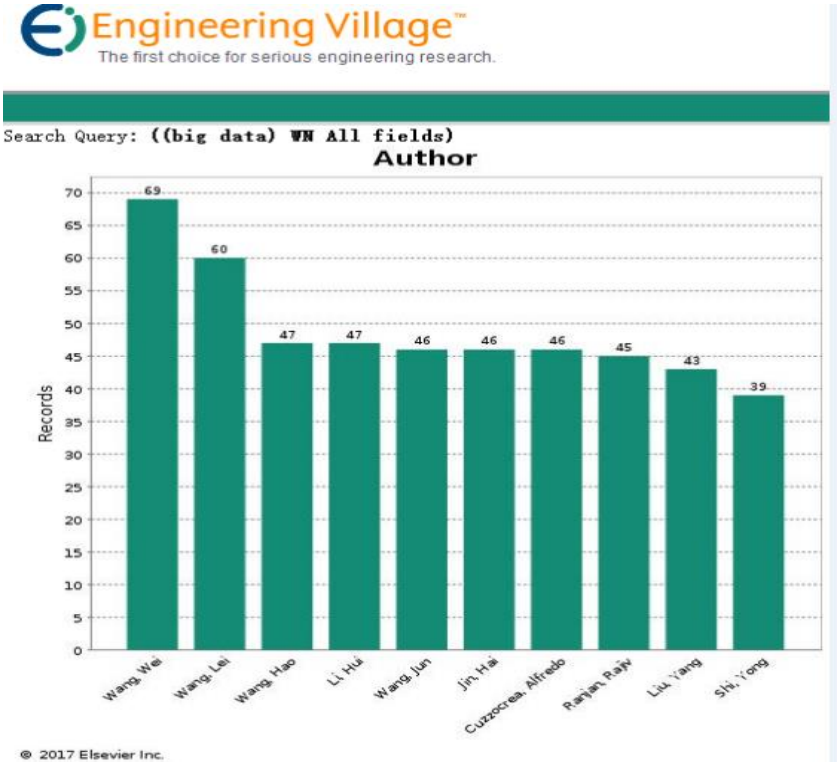

Fig. 1 The authors of researching Big data from greatest to least

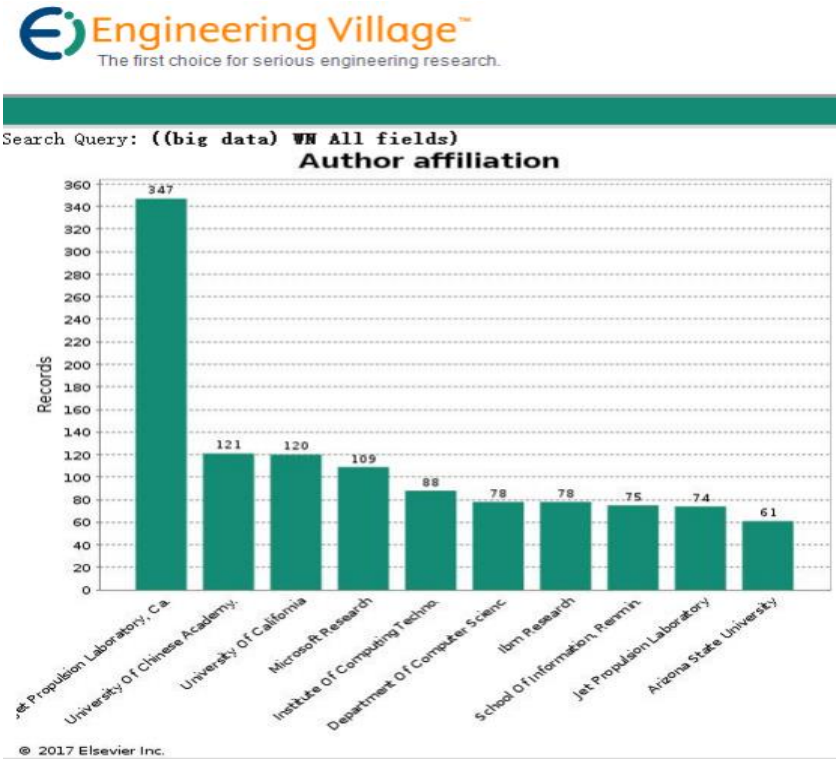

Fig. 2 The authors' affiliation of researching Big data from greatest to least

2. The characteristics of Big data

Viktor meyer-schon berg and Kenneth cook wrote in “The era of Big data”, “Big data refers to not use such shortcut-random analysis method (sampling survey), but adopt the methods of all data [2]), and has 4V characteristics: Volume, Variety, Velocity, and Value”. Subsequently, we explain them in detail as follows.

2.1 Volume

The starting unit of measurement about Big data is at least PB (1024TB), EB (1024 PB) and ZB (1024 EB). For example, every character of two bytes, i.e., 1 character equals 16 bit (2 * 8 bit), i.e., 2 bytes. As we all known that “A dream of red mansions” contains 870000 words, including punctuation, but 853509 words (except punctuation). Therefore, 1GB approximates 671 units of “A dream of red mansions”, 1TB approaches to 631903 units, and 1PB is approximately equal to 647,068,911 units.

2.2 Variety

In this room of Big data, live all kinds of "human", respectively called log, audio, video, chatting records, demographic census, weather forecasts, picture, graphical location information, and so on, multi-type datum put forward higher requirements of the data processing ability.

2.3 Velocity

This is the most notable features of Big data to distinguish the traditional data mining. Existing technical architecture and route has already not efficiently processed such a vast amounts of datum, while for the related organizations, if to put into a huge collection of information cannot get through timely treatment to feedback effective information, it will be a Pyrrhic victory. It can be said that the era of Big data comes up with new challenges to the data control ability of human beings, and also provides an unprecedented space and potential to help us get more profound and comprehensive insight into ability [2]. Take to storage 1 PB datum as an example, even if the bandwidth (internet speed) can reach 1 G/s, the capacity of the computer is enough, and the computer runs 24 hours a day. To store 1 PB datum in the computer also needs 12 days.

2.4 Value

With the wide application of Internet of things (IoT), information awareness is ubiquitous, information is volume, but its value density is relatively lower. How to more quickly complete the value of the data "purification" through powerful machine algorithms, is to be solved problem of the era of Big data.

3. Application examples

"Big data" has been time in areas such as biology, physics, environmental ecology, as well as military, finance, communication, media, and other industries, but which, because of the development of the Internet and information industry, has been given rise to attention in recent years. There are several application examples in the following.

3.1 The politics of Big data

In the process of preparation, data analysis team behind Obama has been in the collection, storage and analysis of voters data. In this election, Obama campaign's senior aides decided to refer to the data analysis results obtained from the team to develop the next campaign. Using the voters in the race for the available actions, behavior, and support of mass datum. For example, on the east coast to find a group of women have the same appeal of celebrities, to copy "clooney effect" and then to raise money for Mr Obama. “The political index of Twitter” provides a measure of users with social media platform for how to evaluate the manner of candidates. Obama's positive emotions index is 59, and Mr Romney’ only 53.

3.2 The public security of Big data

The underlying technology of data mining technology was for research and development of the British six military intelligence to use to track the terrorists at the earliest. The concept of China's big data actually originated from the ministry of public security to grasp Falun gong at the soonest. Big data screens criminal gangs, and two persons with locking criminals take the same train, and lives in the same hotel, which may be accomplices, and in the past, the police officers needed to prove this, by putting together different clues to try to identify the suspect. Through analyzing more and more data mining, certain area of crime and crime patterns will be clearly visible. Big data can help the police locate the most susceptible to the area of lawless intrusion, create a high-crime areas heat maps and schedules, which is not only advantageous for polices to precisely allocate police strength and prevent from fighting crimes, also can help people understand the situation, and be on the alert [3].

3.3 The finance of Big data

Ngari "hydrological model" is respective statistics of a department of Ngari merchants "hydrological data" library by the small micro enterprise category, and level, etc. Like each to the past at some point, the sales of the shop will enter the busy season, and the sales will increase. At the same time each in this period of time, the customer will rise in external invested limit, in combination with the hydrological data, the system can determine the financing needs of stores, and combining with data from the past as funds by the shop and money with data of the similar stores, can judge the money demand amount of this shop [5].

3.4 Effective application of electronic power and Internet of Big data

Effective application of electronic power of Big data has brought the sweeping change. Efficient utilization of electronic power of Big data can provide value-added services to a large number of high value-added business for industry both inside and outside inside, which has a high value for the promotion in level of the electric power enterprise profit and control. The grid experts analyzed, when the data utilization rate turn up 10%, it can make the power grid enhance the profits of 20% ~ 49%. According to the research and analysis of GTM Research, Research and analysis, the market of power Big data management system will reach the scale of $3.8 billion all over the world by 2020, and the power Big data of collection, management, analysis, and the service industry will usher in unprecedented opportunities for development. Meanwhile, the era of Big data applied in the field of Internet has brought a good result, for instance, “Internet adding to tracing” helps more than 500 family return reunion [6].

3.5 The energy source of Big data

Big international oil companies have always placed great emphasis on data management. Such as Chevron connects 50000 desktop systems to 1800 company site, eliminates the refining, marketing and transportation "downstream system" of repeated processes and systems, saves $50 million a year, and for the past four years has won the net present value of about $200 million in return. Accurately predicting solar and wind power needs to analyze large amounts of datum, including wind speed, clouds and other meteorological datum. The manufacturer of Danish Wind turbine, i.e., Vestas Wind Systems, which has been solved by the deployment of the world's largest super computer on IBM's Big data, through the analysis of structured and unstructured volume datum, including the PB level of the weather report\tidal phase, geographical space, satellite imagery, etc., optimizes the Wind turbine layout, effectively improves the performance of Wind turbines, to provide customers with the configuration scheme of precision and optimized Wind turbine that not only helps customers to reduce the cost of per KWH, and enhance the return on investment estimation accuracy of a customer, at the same time, shortens the response time of a business users requesting from a few weeks to a few hours [7,8].

4. Conclusion

In a word, Big data is changing the people's life. Based on the analysis platform of Big data, Governments at all levels, departments, the listed company, the enterprise group, and foreign companies will optimize their decision. The analysis capacity of Big data gradually strengthens, and the traditional market research industry, securities research institute, industrial chain consultancy will disappear step by step. Bank will carry out bank direct selling business based on the enterprise platform of Big data, meanwhile, according to the industrial chain of financial services business model to do business. Because of the emergence of Big data systems, all depending on the information asymmetry of profitable business will disappear. In brief, for government, financial institutions, and enterprises, Big data likes air is indispensable.

Of course, Big data faces many security threats, such as the infrastructure, storage, and data access of security threat, the problem of privacy disclosure, advanced persistent attacks aiming at Big data, and so on. Therefore, we will continue efforts to tackle these issues, and make big data bring us more and more benefits.

Acknowledgements

This work was supported by the Collaborative Fund Project of Science and Technology Agency in Guizhou Province Marked by the word LH on 7487[2014], and the reform project of teaching contents and curriculum system in colleges and universities of Guizhou Province on 2014SJJGXM003, and the project of education and cooperation for talent team word in Guizhou in 2015(NO:[2015]67), the innovative and entrepreneurial training plan of the undergraduates in Guizhou Province in 2014(NO.201410665010), and partly supported by the National Natural Science Foundation of China(NO.61562703).

References

[1] Wu, Lengdong, Yuan, Liyan, You, Jiahuai. Survey of large-Scale data management systems for big data applications, Journal of Computer Science and Technology, vol. 30 (2015), 163-183.
[2] Roy. Ankita, Ray. Soumya, Goswami, Radha Tamal. Approaches and challenges of big data analytics-study of a beginner, Advances in Intelligent Systems and Computing, vol. 458 (2017), 237-245.
[3] H. Bo, T. Fei, X. Jianfeng. A survey on job scheduling in big data, Cybernetics and Information Technologies, vol. 16 (2016), 35-51.
[4] Y. Wenquan, T. Fei, X. Jianfeng, Noval stream data mining framework under the background of big data, Cybernetics and Information Technologies, vol. 16 (2016), 69-77.
[5] M. Bharat B., B. Manoj. Attack tolerant architecture for big data file systems, Performance Evaluation Review, vol. 41 (2014), 65-69.
[6] Y. Yindong, B. Zhongqin, Advances and future challenges in electric power big data. Proceedings-2014 2nd International Conference on Advanced Cloud and Big Data, CBD 2014, August 3 (2015), 213-219.
[7] G. Priya, S. Nuzhaft, Improving network traffic in Map Reduce for big data applications, International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016, November 22 (2016), 2979-2983.
[8] B. David, K. T. Dunn, M. Bill, Big data, big data quality problem, Proceedings-2015 IEEE International Conference on Big Data, IEEE Big Data 2015, December 22 (2015), 2644-2653.