

Big data analytics and the problems with user security information processing

Baoxin Zhang*

School of Computer and Information Technology, Beijing Jiao tong University, Beijing 100044, China

*Corresponding author: 19722074@bjtu.edu.cn

Keywords: Big data analysis, Information processing, User security, Storage mechanism.

Abstract: While big data analytics may be deemed necessary for modern society and its progress, this necessity comes with some significant concerns and obstacles. Big data analytics is defined by its ability to cope with massive amounts of data and the sharing of that data. This raises a number of security concerns, particularly when working with data that crosses sectors and businesses. The term "big data" refers to digital data stores with a high volume, velocity, and diversity; big data analytics is the process of utilizing software to discover trends, patterns, correlations, and other important insights inside those massive data stores (Gupta and George, 2016). It also plays a critical role in the creation and development of numerous industries, although current constraints restrict its development to a great extent. As a result, resolving the issues impeding its development takes precedence. The aim of this paper is to discuss the impact of big data analytics on mankind, to discuss the problems encountered in its development and the possible solutions to these problems. First, it will enumerate the application of big data analytics in many areas, and demonstrate its widespread influence and the critical need to enhance big data analytics' ability to contribute to societal development. Then the development challenges and effects will be analysed and finally the potential solutions will be assessed and discussed.

1. Introduction

'Information is the oil of the 21st century, and analytics is the combustion engine'. So said Peter Sondergaard in 2011 and what may have been true then certainly seems true today as the development of big data. Seen as one of the most valuable commodities in managing automation systems because of its availability of data from different sources, big data is defined as "information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization (Beyer and Laney, 2012)". Although the information could be collected and integrated more easily in the global big data network which could bring a decrease in capital costs and an increase in corporate profits, the massive database brought by big data makes the users' information security inadapttable.

2. The application of big data

The rapid development of the Internet catapulted humanity into an era of tremendous information expansion, and people's lives began to shift online as well. With the globalization of people's lives, the amount of information in the whole society has exploded, and "big data" has become the inevitable product of the "data" trend (Gupta and George, 2016). Many enterprises still use their prior personal experience in decision-making rather than on data analysis, which will lead to problems and complications during the actual decision-making operation. In the mature countries of data analysis industry, 90% of market and business decisions are determined by big data analytics. The widespread use of big data analysis can assist enterprises in comprehending additional elements and integrating into the market. Utilizing data and paying attention to quantitative analysis has evolved into a critical component of scientific research, company management, and government decision-making.

Not only are the economic benefits of big data analysis significant; the fast growing field of big data analysis has begun to play a critical role in the development of health care practice and research.

It gives instruments for analyzing and absorbing a significant volume of wholly disparate data created by the existing health care system in contemporary medical research. Using a large number of medical data combined with multi-mode data from different sources for the latest research (Belle, et al., 2015). Additionally, big data has made a significant contribution to social security. Criminal activity has always posed a significant threat to societal security, and combating crime has been a critical duty of all countries and areas, and has never ceased to be (Akhgar, et al., 2015). It may integrate physical analysis, chemical analysis, medicine, and other professional expertise to analyses the crime scene and form reasonable inferences in the action of combating crime. When interrogating detainees, they can use criminal psychology and linguistics to analyses the detainees' thoughts in order to assist police in making accurate conclusions. It carries on the record research to the railway in terms of transportation, by comparing the railway track in the database to the present research track, and conducts risk assessment, and this can significantly improve passenger safety.

3. Problems faced by big data analysis

As the most active component of technological innovation, big data is reconstructing the global economy's production, distribution, consumption and other fields. The reconstruction has a comprehensive and profound impact on global competition, national governance, economic development, social life. All of these factors contribute to the overall betterment of society's level of living. As a result, developing big data analytics is a collective responsibility; it is an international endeavour that requires each country to contribute its own resources.

Currently, the problem of big data analysis is primarily separated into two dimensions. One is its own development problem, the science and technology problems, in the development of various technical problems restricted its role, and the root of these problems is information processing (Gupta and George, 2016). The term "information processing" encompasses a broad range of activities, including data storage, resource allocation, and information sharing. The second main problem is user security. The generation of big data makes data analysis and application more complex and difficult to manage, highlighting the prominence of data security and privacy protection.

3.1 Information processing

If the traditional data storage method is retained in information processing, it will create a slew of inconvenience for big data analysis. At the time point of big data generation, the volume of data generated is enormous and difficult to quantify. However, in the early stages, the architecture designed for traditional storage mode is only suitable for a small amount of data, and the amount of data currently stored is increasing rapidly, pushing the architecture's capability to its limit and potentially causing the large amount of data to change and lose integrity (Jun, et al., 2016). As a result, the massive volume of data cannot be directly stored by the traditional structured database. The storage of data is the foundation for big data analysis. Resource allocation and scheduling are needed after storage. Because resources are limited, allocating scarce resources to various computer workloads has become a critical issue. If resource allocation is unreasonable, it may result in an emergency going unresolved. Information sharing among departments is critical for improving the efficiency of data analysis. However, data sharing between departments and industries remains limited, resulting in what is known as an information isolated island. And it has resulted in a slew of difficulties, including inefficient labour, resource waste, and a loss of benefits to businesses. The existence of information isolated island prevents data from forming valuable comprehensive information, which leads to numerous departments and industries in the field of research lacking reference basis, affecting the application of the whole information industry.

3.2 Users security

With the development of e-commerce and mobile phone online behaviour, users will leave personal information on the Internet. The huge amount of information is stored in a distributed storage mode. The storage path view of this storage mode is relatively clear, and hackers can readily exploit relevant

weaknesses to conduct illicit actions. These criminals will benefit from this data by selling the user's personal information, defrauding you or your family, and severely injuring the user's personal rights and interests (Hasan, et al., 2013). The information security problems of this kind of big data analysis seriously affect the data integrity and the performance of big data system. It also severely affects the safety of the public.

4. Dynamic storage mechanism for alleviating the problems

It is critical to resolve these issues in order to improve the utility of big data analysis for humans. To tackle the information processing problem, firstly, it needs to figure out the weakness in data storage. The traditional storage mode is static and cannot adapt to the uncertainty of big data, it is necessary to explore a data storage mode suitable for big data, data dynamic processing technology can be used. Building dynamic storage mechanism is to deal with the cause. Dynamic mechanism can handle with the characteristics of data variability. Next, looking at the resource scheduling problem, A dynamic response mechanism is needed to assign priority to each computer job, and will be sorted according to these values, giving priority to resource scheduling for those urgent events. Reasonable configuration and scheduling of limited computing and storage resources. It deals with the consequence, working out the problem that important items cannot be dealt with because of limited resources. Last, the information sharing problem. It is necessary to combine the information in different fields and build a platform for data sharing, so as to give full play to the advantages of big data, and make all kinds of industries have data to quote when they are taking a research not belonging to their area. It deals with the cause, after the establishment of a sharing platform, the data sharing before different departments can be strengthened, and the problem of data shortage can be disposed to a large extent.

There are two key factors to mitigate the user security problem. The first is from the administration; the government should enhance the corresponding rules and regulations governing user safety, stiffen the penalties, and harshly penalize users who violate their privacy. It deals with the problem with the consequence, and its purpose is to let someone like hackers does not steal others' sensitive data. The second comes from technical means, such as data combing and classification. Aiming at the core data, a series of protective operations such as risk assessment and leakage warning blocking are carried out, and displayed in a visual form to achieve the effect and purpose of effectively ensuring the data. It handles with the cause, protecting user data through technology.

The establishment of dynamic storage mechanism can greatly relieve the storage pressure. It is allocated by the system according to the application's requirements and does not need the software to pre-allocate storage space like the static memory allocation approach does. There is also no waste of storage space when it is not used. However, for the current dynamic storage mechanism, there is also a few problems existing, the dynamic storage allocation algorithm has the disadvantages of low efficiency, high space and time overhead, and the development cost is much higher than that of static storage. If the dynamic storage mechanism is capable of effectively lowering development costs and increasing space usage efficiency, it must be a viable solution to the storage problem. The purpose of the dynamic response mechanism is to prioritize computer operations, allowing for more efficient resource scheduling and optimum utilization of scarce resources (Wang, et al., 2011). This has a great effect on solving the problem of resource scheduling in the short term, but it is not a long-term solution. Although priority has been given to urgent tasks, in the end there are still many unresolved problems due to limited resources, and in the future we should find a fundamental solution to the problem. A shared platform is established to alleviate the problem of information isolation between different departments, this is a feasible solution, and it can collect all kinds of information and provide convenience for people to find corresponding data. Through the improvement of the above three steps of information processing, the problem of information processing will be settled to a great extent.

5. Conclusion

In a conclusion, big data analytics brings convenience to social life, but also has many problems. Although these problems will have adverse effects on society, but any technology has no absolute advantages and disadvantages, it should be treated dialectically. By optimizing data storage and resource scheduling, as well as by establishing a shared platform, the problem of information processing in big data analysis can be resolved. By combining and classifying data via technical means, we can address the issue of data that is easily attacked, while simultaneously enhancing the corresponding rules and regulations to protect users' privacy from a government perspective. They are both attempting to promote big data analytics; their objective is to identify issues with big data analytics and to maximize its benefits. There is every reason to believe that this may create a mature big data analytics system that benefits the entire mankind.

References

- [1] Akhgar, B., Saathoff, G., Arabnia, H., Hill, R., Staniforth, A. and Bayerl, P. (2015) Application of Big Data for National Security: A Practitioner's Guide to Emerging Technologies. *Big data for national security*, 32(8), 1345-1365.
- [2] Belle, A., Thiagarajan, R., Soroushmehr, S., Navidi, F., Beard, D., Najarian, K. and Li, X. (2015) Big Data Analytics in Healthcare. *BioMed research international*, 2015(4), 1253-1269.
- [3] Beyer, M.A. and Laney, D. (2012) the Importance of "Big Data": A Definition. Gartner, G00235055.
- [4] Gupta, M. and George, J. (2016) toward the development of a big data analytics capability. *Information & management*, 53(8), 1049-1064.
- [5] Hasan, O., Habegger, B., Brunie, L., Bennani, N. and Damiani, E. (2013) A Discussion of Privacy Challenges in User Profiling with Big Data Techniques: The EEXCESS Use Case. 2013 IEEE International Congress on Big Data, 25-30.
- [6] Jun, S., Liu, M., Lee, S., Hicks, J., Ankcorn, J., King, M., Xu, S. and Arvind (2016) BlueDBM: Distributed Flash Storage for Big Data Analytics. *ACM Transactions on Computer Systems*, 34(3), 1-31.
- [7] Wang, Q., Wang, C., Ren, K., Lou, K.J. and Li, J. (2011) Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing. *IEEE Transactions on Parallel and Distributed Systems*, 22(5), 847-8.