Data Privacy and Security Challenges in Big Data Analytics: A Review of Current Solutions and Future Directions

Muhammad Khuram Khalil[1], Asala Ahmed Sulaiman Al Amri[1], Miad Khalid Abdul Latif Al Balushi*, Suad Al Qassabi[1#] and Vikas Rao Naidu[1#]

[1]Middle East College, Muscat, Oman
*Translator
#Advisor

ABSTRACT

The growing amount of data generated by our daily activities, known as big data, presents distinctive challenges due to its scale, complexity, and velocity. Big data analytics is the process of collecting and analyzing this data to uncover insights and gain a competitive advantage. However, as the use of big data continues to increase, the importance of data privacy and security has become more critical. The sheer size and complexity of big data sets can make it difficult to implement privacy and security algorithms that work at the necessary scale and speed. Therefore, there is a need to explore current solutions and future directions for addressing the data privacy and security challenges in big data analytics. This review aims to provide an overview of the current state of knowledge in this field, including a discussion of the challenges and potential solutions for ensuring data privacy and security in the context of big data analytics. By identifying the gaps in the literature and proposing future research directions, this review intends to contribute to a deeper understanding of the data privacy and security challenges in big data analytics and inform the development of more effective solutions.

Introduction

With the increment of the use of big data, it becomes important to focus on enhancing the privacy and security of data. Big data refers to extremely large and complex data sets that are difficult to manage, process, and analyze using traditional data processing methods. Identifying patterns, insights, and trends in large and complex data sets is the function of big data analytics. This involves using advanced analytical techniques such as machine learning, data mining, and predictive modeling to extract valuable insights and inform decision-making. Big data analytics helps organizations to understand their customers, optimize their operations, and gain a competitive advantage in their industry. Developing a system that can handle the variations, speed, and size of big data is essential for protecting sensitive data and ensuring the correctness and integrity of information. The system may face a lot of risks like stealing data. And it will be more dangerous especially when they store sensitive data like credit card information and user confidential information (Antonenko, 2022).

Big Data Security and Privacy Challenges

Big data has a lot of effects on data security and privacy so, they become challenges big data could avoid or could not. There are some of the challenges that big data security and privacy may face:

- Access control

Organizations restrict access to the confidential data of users like user medical data which includes personal details. But some people need to access this data other than doctors who are medical researchers, and they don’t have permission to access the data. To solve this problem most organizations give them specific access which is they can access only the data they need it. The design of big data is not developed for specific access so, the solution is to copy all medical data without personal information about the patient.

- The fast development of NoSQL databases and less focus on security considerations
  Most organizations used NoSQL databases which are popular in big data. And because it is popular that will cause a lot of security problems. They are developing the NoSQL database constantly with new technical features. But they ignore security, which is more important than other features.
- Data storage
  Most organizations used the cloud to increase the speed of data transfer and its operations. But security risk in the cloud is increasing because anyone can have access to important data even with the smallest supervision in the access of data control. Avoid big organizations, storing important data in the workplace but fewer sensitive data are stored in the cloud to obtain security. But to add policies to the database cybersecurity is needed to access important data even with the smallest supervision in data control access which increases the cost of the database. However, it is better to pay more to avoid big data risks.

There are big a lot of challenges for big data, but the smallest mistake made big problems. However, they need to develop the system carefully focusing on every detail.

Literature Review

Data Security and Privacy

Every user used privacy and security in every program or organization because these two things are the most important things for the user. Privacy has control over your data and how it can be viewed and used. Security protects your data from unauthorized access or any damage your data may face (Privacy vs. Security: Exploring the differences & relationship, 2022) In the article, the author proposes a chance in the future that privacy and security will be merged and developed to have a better version of them. In my view, privacy, and security should be merged from the beginning because they will complete each other they are all about important user data and how to protect it in diverse ways in privacy they do not let anyone enter your data which that help to protect the data. On the other hand, security protects your data and does not let anyone access it. However, merging privacy and security will have a lot of benefits for the users and people who will develop the system of privacy and security because in the future it will be one of the major trends that all companies want to add to their system.

Big Data Security and Privacy: Current Challenges and Future Research perspective in Cloud Environment

The article discussed some of the current challenges and future direction views that are related to big data security and privacy in the cloud environment. This article focuses on the importance of the security and privacy of data in the cloud. And it discussed the challenges and risks that we may face if we store a large amount of data in the cloud. Moreover, the challenges are related to security data issues and cybercrime on sensitive data. Cloud computing is an important aspect of our devices because it stores all our data. Additionally, it is used for organizations to store all sensitive data. To solve these challenges and risks the authors suggest solutions such as encryption techniques, controlling the access to sensitive data, how to improve the system of security and privacy to make it stronger than the current time, and ensuring that security and privacy system need to be improved through continual development.

Overall, this article takes a general look at big data security and privacy in the cloud environment. This article provides useful information about the importance of solving these risks for our sensitive data security and privacy, and how to secure our sensitive data from cybercrime. Moreover, these articles give us many solutions to solve these challenges.

Big Data Privacy: A Technological Perspective and Review

The authors of this article discuss the security and privacy challenges of big data in Internet of Things (IoT) environments. They highlight the growing importance of big data security and privacy in the IoT environment. Several key challenges and risks are identified in the article along with collecting, storing, and processing large amounts of data. These challenges include issues related to data confidentiality, integrity, and availability, as well as concerns about the misuse of sensitive data and the potential for cyberattacks. In order to address these challenges, the authors propose several possible solutions, including the use of encryption techniques, data anonymization, and privacy-preserving data analysis. Furthermore, they emphasize the necessity of developing robust policies and protocols for security and privacy, as well as an ongoing research and development effort. Authors further argue that collaboration among different stakeholders, including researchers, practitioners, and policymakers, is essential for effectively addressing the security and privacy challenges of big data in IoT environments. As well as the authors suggested that the standard or best practice regarding big data security and privacy could also lead to greater interoperability across different systems and applications in IoT. Overall, this article provides a useful overview of the current state of research on big data security and privacy in IoT environments. It also focuses on the importance of addressing these issues in order to ensure the privacy, integrity and availability of data, as well as to protect sensitive data from malicious destruction and cyberattacks. The article finally identifies a number of potential solutions and strategies for addressing these challenges (Jain et al., 2016).

Methods

The purpose of this section is to provide an overview of the tools, methods, and procedures used to collect data for this study. Research design: The research is a systematic literature review that synthesizes existing research on data privacy and security challenges in big data analytics. The review is examining current solutions and future directions for addressing these challenges. Search strategy: The search strategy involves searching relevant databases such as ResearchGate, and Google Scholar. The search terms include "data privacy," "data security," "big data analytics," "challenges," "solutions," and "future directions." The search is limited to articles published in the last 10 years.

Conclusion

In conclusion, the importance of data privacy and security in big data analytics has become increasingly crucial, given the ever-increasing volume and complexity of data. The challenges associated with securing big data are also growing, making it essential to explore current solutions and future directions for mitigating these challenges. This review has provided a comprehensive overview of the importance, challenges, current solutions, and future directions of data privacy and security in big data analytics. By synthesizing the existing literature, the review has highlighted the need for more effective security measures to protect against cybercrime and data misuse. Overall, the research has contributed to a deeper understanding of the data privacy and security challenges in big data analytics and has identified potential solutions and future research directions. By increasing awareness of the importance of data privacy and security, this review aims to help individuals and organizations protect themselves against the risks associated with big data analytics.

References

Antonenko, D. (2022, March 14). *Big Data Privacy and Security Challenges: What you need to know.* Businesstechweekly.com. https://www.businesstechweekly.com/operational-efficiency/data-management/big-data-privacy-and-security-challenges/

Jain, P., Gyanchandani, M., & Khare, N. (2016). Big data privacy: a technological perspective and review. *Journal of Big Data*, *3*(1). https://doi.org/10.1186/s40537-016-0059-y

Nesvit, A. (2022, December 7). *What is the future of data privacy and data protection.* Infopulse SCM. https://compliance-aspekte.de/en/articles/main-data-privacy-trends-to-watch-in-2022-2025/

Privacy vs. Security: *Exploring the differences & relationship*. (2022). Okta.com, https://www.okta.com/identity-101/privacy-vs-security/

Zia, T., Shah, M. A., Abbas, H., Qaisar, S., & Zomaya, A. Y. (2020). *Big Data Security and Privacy: Current Challenges and Future Research perspective in Cloud Environment.* Journal of Parallel and Distributed Computing, 146, 28-49. doi: 10.1016/j.jpdc.2020.05.003