IGNITE : Journal Islamic Global Network for Information Technology and Entrepreneurship
Volume. 2, No.1 Januari 2024
e-ISSN :3025-2857; dan, p-ISSN :3025-2849, Hal 49-60
DOI: https://doi.org/10.59841/ignite.v1i4.

Development Of Data Security Algorithms: A Literature Review On Information Security In The Context Of Big Data

Nathanael David Christian Barus[1]
China Three Gorges University, Yichang 443002, People's Republic of China*;
Email : nathanaeldavid.idn@gmail.com

Natasha Fedora Barus[2]
Telkom University, Bandung 40257, Indonesia
Email : natashafedorab@gmail.com

Abstract: *In the era of Big Data, securing sensitive information and ensuring data integrity have become paramount concerns due to the unprecedented volume and intricacy of data. Traditional security algorithms face significant challenges in adapting to the distinct characteristics of Big Data. This literature review explores the evolution of data security algorithms tailored explicitly for the Big Data landscape, aiming to address the increasing demand for robust security solutions capable of handling the unique challenges posed by the massive scale and complexity of data. By scrutinizing existing literature, the review unveils advancements, trends, and innovations developed by researchers and practitioners to mitigate vulnerabilities associated with handling vast datasets. The review also sheds light on emerging technologies and cryptographic techniques specifically designed for Big Data security, contributing to enhanced confidentiality, integrity, and availability in the face of evolving cyber threats. While these developments offer advantages such as improved data protection and threat detection, the review highlights challenges, including algorithmic bias, computational complexity, privacy trade-offs, and a shortage of skilled workforce. By considering these factors and emphasizing continuous improvement and ethical considerations, organizations can responsibly leverage data security algorithms to enhance information security in the era of Big Data.*

Keywords: *Big Data, Security, Algorithm, Literatur Review.*

INTRODUCTION

The explosion of information in the digital era not only brings ease of access but also poses new challenges, especially regarding the security of large-scale data known as Big Data. Organizations across various sectors continue to generate and collect data in massive volumes, creating opportunities for valuable insights while also posing vulnerabilities to sensitive information (Li et al., 2018). Data security is no longer just an issue for institutions; it has become a fundamental concern for individuals and even governments. In the context of Big Data, scholarly literature serves as a lens to explore the landscape of data security algorithms. This literature review aims to provide a deeper understanding of the challenges in securing Big Data, examining algorithms that act as defense mechanisms, and reviewing recent developments in this domain (Dal Pozzolo et al., 2014).

Security in Big Data has become a focus of attention for several crucial reasons. The ongoing influx of data volume surpasses the capacity of conventional technology and security methods. In this context, the principle of "Think big, secure bigger!" becomes essential to confront the escalating scale. Additionally, the diversity of data types within Big Data is a key

Received: Desember 29, 2023; Accepted: Januari 09, 2024; Published: Januari 09, 2024
* Nathanael David Christian Barus , nathanaeldavid.idn@gmail.com

factor. Beyond numbers and text, Big Data encompasses a wide spectrum, including semi-structured and unstructured data such as images, videos, and digital social media traces (Shaukat et al., 2020). This complexity requires a more sophisticated security approach. The openness to extensive analytics in Big Data, involving collaboration and data merging from various sources, can be a potential gap for leaking sensitive information. Therefore, security measures must be taken with extra caution to address this risk (Sharafaldin et al., 2018). Furthermore, the evolution of increasingly sophisticated cyber attacks adds urgency to the need for adaptive and evolving security algorithms. Cybercriminals continue to develop new methods to breach and damage data, making Big Data security always ready to face evolving threats. In this context, the exploration of security algorithms is key to maintaining the integrity and security of data in the era of Big Data (Gheyas & Abdallah, 2016).

Big data processing opens up new opportunities through its analytical capabilities, providing valuable insights across various business sectors. Industries such as automotive, energy distribution, healthcare, and retail can leverage big data for better decision-making. For instance, analyzing driving patterns can reveal anomalies in behavior, smart network data can facilitate energy load forecasts, search engine queries can help detect influenza outbreaks, and customer purchase history can be utilized for personalized recommendations (Moghadam & Colomo-Palacios, 2018). However, the common thread among these applications is the inclusion of individual-related data, introducing sensitivity aspects. Despite its analytical benefits, big data faces challenges in efficient storage, management, and processing due to its intrinsic characteristics. Initially classified into three dimensions as the three Vs (volume, variety, and velocity), this framework has evolved to include newer Vs such as veracity and value. Volume, referring to the large amount of data, is considered significant when exceeding 10^18 bytes (Rawat et al., 2021). Variety acknowledges diverse data formats, including text, numbers, video, and images. Velocity reflects how quickly new data is generated. Veracity emphasizes the importance of accurate and reliable data, while value highlights the usefulness of specific data points or combinations. With the potential for large-scale data processing, there is a demand for efficient and scalable solutions that prioritize security and privacy (Georgiadis & Poels, 2022).

As the framework of the five Vs increasingly dominates, addressing the challenges of volume, variety, velocity, veracity, and value becomes crucial to unlock the full potential of big data in various industries. This literature review details various algorithms that play a frontline role in ensuring the security of Big Data. First, there is encryption, an invisible cloak that makes data unreadable except by those with a specific "key." This method provides extra protection by securing the integrity and confidentiality of data. Next, data masking presents another approach by obscuring sensitive information behind a "mask," preserving its analytical value without compromising it (Singh et al., 2018). Access control acts as a barrier gate that regulates who can access data and the limitations imposed on that user. The journey to secure Big Data continues to evolve, and this literature review outlines trends and recent research. Homomorphic-based security, an innovative concept, allows direct computation on encrypted data without the need for decryption first, preserving privacy without sacrificing functionality (Chandra et al., 2017). The use of blockchain technology for Big Data security provides a solution to ensure data integrity and authenticity through decentralized blockchains. Meanwhile, machine learning for anomaly detection uses artificial intelligence to identify suspicious activities that may indicate cyberattacks.

By understanding the dynamics of data security algorithms and recognizing the challenges faced by Big Data, recommendations can be formulated to build a more robust and reliable system. One key step is selecting the right combination of algorithms, considering that there is no single solution suitable for all cases (Lei Xu et al., 2014). Continuous updates to security algorithms and techniques are also essential in response to the evolving landscape of cyber threats. Enhanced security awareness through user and personnel education is a crucial pillar in shaping a strong security culture. Additionally, cross-sector collaboration with experts and stakeholders is necessary to develop more comprehensive security solutions (Ismagilova et al., 2022). Through this literature review, not only is the importance of Big Data security understood, but insights into applicable algorithms and approaches are also gained. By combining these various security shields, a secure, reliable, and potentially progressive future for Big Data can be collectively created across different fields.

RESEARCH METHOD

This research will adopt a qualitative methodology with a focus on literature review to develop data security algorithms in the context of big data. The literature review approach will serve as the primary framework for collecting and analyzing information related to information security in the big data environment. The first step in this research methodology involves identifying relevant literature sources on the topic of developing data security algorithms. These literature sources may include scholarly journals, books, conferences, and other related articles that discuss information security aspects in the context of big data. After identifying the literature sources, the next step is to conduct a thorough review of these literatures. This will include an in-depth analysis of recent developments in data security algorithms, encryption techniques applicable in the big data environment, and other relevant approaches to protect sensitive information in big data.

In the next stage, the research will evaluate the strengths and weaknesses of the identified security algorithms, considering security aspects such as data integrity, confidentiality, and availability. This analysis will provide deep insights into the effectiveness of these algorithms in safeguarding information security in big data scenarios. Finally, the research will draw conclusions and recommendations based on the findings from the literature review, identifying potential directions for future research and highlighting key aspects to be considered in the development of data security algorithms in the era of big data. Using a qualitative literature review approach, this research aims to make a valuable contribution to the understanding and further development of information security in the context of big data.

RESULT AND DISCUSSION

In the era of Big Data, where the sheer volume and intricacy of data have reached unprecedented levels, safeguarding sensitive information and ensuring data integrity have become paramount concerns. The conventional security algorithms, which have long been effective in traditional data settings, encounter significant challenges when confronted with the distinct characteristics of Big Data. This literature review embarks on an exploration of the evolution of data security algorithms tailored explicitly for the Big Data landscape (Cremer et al., 2022). The objective is to address the escalating demand for more robust security solutions that can effectively grapple with the unique challenges posed by the massive scale and complexity of data in contemporary environments. The review aims to provide a comprehensive understanding of the current state of information security within the realm of Big Data. By scrutinizing existing literature, it seeks to uncover the advancements, trends, and innovations that researchers and practitioners have developed to mitigate the vulnerabilities associated with handling vast datasets (Mvula et al., 2023). This involves an in-depth analysis of novel approaches and methodologies that go beyond the limitations of traditional security measures, acknowledging the need for adaptive and scalable solutions.

Furthermore, the literature review is likely to shed light on emerging technologies and cryptographic techniques that have been specifically tailored to address the intricacies of Big Data security. It explores how these advancements contribute to enhancing the confidentiality, integrity, and availability of information in the face of evolving cyber threats. Through a critical examination of the literature, the review aims to offer insights into the ongoing efforts to establish a robust foundation for information security in the dynamic landscape of Big Data, paving the way for the development of more effective and resilient data security algorithms.

Key Challenges of Big Data Security

The security landscape in the realm of Big Data is fraught with unique challenges that stem from the sheer volume and diversity of data. One prominent challenge lies in the volume and variety of data, where traditional algorithms struggle to efficiently process massive datasets. This inefficiency not only leads to performance bottlenecks but also opens avenues for potential security vulnerabilities. As the scale of data continues to grow exponentially, finding solutions that can handle this data deluge becomes imperative to maintain robust security (Sumithra & Parameswari, 2022). Heterogeneity is another key challenge, as Big Data often encompasses diverse data formats, including structured, semi-structured, and unstructured data. This necessitates the development of flexible algorithms capable of adapting to different data types, ensuring that security measures are not constrained by the varying formats inherent in large datasets (Li et al., 2018).

Privacy concerns pose a critical challenge in the Big Data landscape. Balancing the extraction of valuable insights from vast datasets with the imperative to protect individual privacy is a delicate task. The challenge lies in devising security solutions that can strike the right balance between utility and anonymity, ensuring that valuable information can be extracted without compromising the privacy of individuals. Distributed storage and processing further compound the security challenges in Big Data environments.

With data distributed across multiple locations, secure communication and processing protocols are essential to prevent unauthorized access. Ensuring the integrity and confidentiality of data throughout its distributed journey becomes a complex task, requiring sophisticated security measures that can adapt to the decentralized nature of Big Data systems. Addressing these key challenges is crucial for the development of effective security frameworks tailored to the unique characteristics of Big Data (Dal Pozzolo et al., 2014). As technology continues to advance, the exploration of innovative solutions becomes paramount to safeguarding the integrity, confidentiality, and privacy of information in the era of massive and diverse datasets.

Big data security is a multifaceted and dynamic field characterized by various challenges that organizations must navigate to ensure the integrity, confidentiality, and accessibility of vast datasets. One of the primary challenges revolves around data storage, given the sheer volume of information involved. Securing this stored data is crucial to prevent unauthorized access, and implementing robust access controls becomes imperative to restrict data access to authorized personnel only (Shaukat et al., 2020). Another significant challenge is the vulnerability of big data to fake data generation. The introduction of inaccurate or fraudulent information into the system can lead to distorted results and misguided decision-making. Detection and prevention mechanisms are vital to ensure the accuracy and reliability of insights derived from big data analytics.

Data privacy is a persistent concern in the realm of big data, particularly as it often involves sensitive information. Encryption and decryption play a critical role in protecting data privacy by securing access to and storage of sensitive information, thereby mitigating the risk of unauthorized access and potential breaches. Efficient data management is a challenge given the complexity and sheer volume of big data (Sharafaldin et al., 2018). Organizations must implement effective data management systems to ensure the accuracy and currency of the data, facilitating proper organization and maintenance.

Controlling access to big data poses a challenge due to the large number of users who may require access. Implementing robust access control mechanisms becomes essential to ensure that only authorized individuals can access and manipulate the data, preventing unauthorized use or tampering. Data poisoning attacks are a significant threat to big data integrity. Malicious actors may introduce false data into the system to manipulate results. Organizations must deploy mechanisms to detect and prevent such attacks, safeguarding the reliability of analytical outcomes.

Employee theft is a considerable concern, necessitating effective security measures to prevent employees from stealing or misusing sensitive data. This involves implementing strict access controls, monitoring user activities, and fostering a culture of data security within the organization (Gheyas & Abdallah, 2016). To address these challenges, organizations can leverage powerful cryptographic algorithms such as RSA, AES, and One Time Pad. Additionally, innovative solutions like the proposed Secure Dynamic Bit Standard (SDBS) algorithm offer enhanced security, emphasizing the importance of continually evolving data security strategies to meet the dynamic landscape of big data threats. Overall, a comprehensive and adaptive approach is crucial to effectively tackle the multifaceted challenges inherent in big data security.

Emerging Data Security Algorithms for Big Data

The rapid evolution of Big Data has prompted the development of innovative data security algorithms designed to tackle the unique challenges posed by massive and diverse datasets. One notable advancement is Homomorphic Encryption, which enables computations on encrypted data without the need for decryption. This preserves the confidentiality of sensitive information while allowing for meaningful analysis, providing a crucial layer of security in data processing workflows. Attribute-Based Encryption represents another frontier in data security for Big Data. This approach grants access based on user attributes rather than specific identities, enhancing access control and providing a more flexible and dynamic security framework (Moghadam & Colomo-Palacios, 2018). This adaptability proves particularly valuable in environments where diverse users may require varying levels of access to different facets of the data.

Federated Learning has emerged as a cutting-edge solution to privacy concerns in Big Data environments. This approach enables collaborative training of machine learning models on decentralized datasets without sharing raw data. By facilitating model updates without centralizing sensitive information, federated learning addresses privacy concerns and ensures that valuable insights can be gleaned without compromising individual data privacy (Rawat et al., 2021). Blockchain-based solutions have gained prominence for securing Big Data. Leveraging distributed ledgers, these solutions provide a robust foundation for secure data storage, access control, and the creation of audit trails. The decentralized and tamper-resistant nature of blockchain technology enhances the overall security posture of Big Data systems.

Additionally, privacy-preserving data mining techniques have become integral in the quest for secure insights. Methods such as differential privacy and secure multi-party computation allow the extraction of valuable information from datasets while anonymizing sensitive data. These techniques play a pivotal role in ensuring that data mining processes do not compromise individual privacy, a critical consideration in the age of heightened privacy concerns and regulations (Singh et al., 2018). In summary, the emergence of these data security algorithms represents a concerted effort to fortify the security infrastructure of Big Data. As technology continues to advance, these innovative solutions contribute to building a more resilient and privacy-conscious framework for handling the challenges posed by the ever-expanding and diverse landscape of Big Data.

The development and implementation of emerging data security algorithms for Big Data play a pivotal role in safeguarding data protection and privacy, particularly in the context of Big Data cloud environments. A recent research paper underscores the significance of Big Data cloud in providing expansive capacity and innovative informatics, enhancing the processes of discovery, decision-making, and overall process improvement (Chandra et al., 2017). The paper highlights the use of encryption and decryption within a cloud computing system, emphasizing their role in securing access to and storage of data. Notably, the authors introduce a novel algorithm called Secure Dynamic Bit Standard (SDBS), designed to offer high-security levels for data stored by end-users. This innovative algorithm represents a proactive step towards addressing the evolving challenges of data security in the dynamic landscape of Big Data.

In a related article, the focus is on protecting cloud data through the application of robust cryptographic algorithms, including RSA, AES (Advanced Encryption Standard), and One Time Pad (OTP). The article emphasizes the use of encryption and decryption as fundamental tools for accessing and storing data securely within cloud environments. By leveraging powerful cryptographic techniques, organizations can fortify their data security measures and protect sensitive information from unauthorized access and potential breaches. These insights from recent research underscore the growing recognition of the critical role played by advanced data security algorithms, cryptographic techniques, and encryption protocols in the realm of Big Data (Lei Xu et al., 2014). As organizations continue to harness the capabilities of cloud computing and grapple with the challenges posed by vast and complex datasets, these emerging algorithms and encryption methodologies serve as integral components of a robust and proactive approach to data protection and privacy in the digital age.

Advantages and disadvantages of Developing Data Security Algorithms for Big Data

The evolution of data security algorithms for Big Data presents a nuanced landscape with both advantages and disadvantages, necessitating careful consideration. On the positive side, these advancements offer enhanced data protection capabilities. Specifically tailored algorithms can effectively grapple with the challenges inherent in securing vast and diverse datasets, mitigating risks associated with unauthorized access, data breaches, and manipulation. The development of algorithms that are responsive to the unique characteristics of Big Data contributes significantly to fortifying the overall security posture of organizations operating in this data-intensive landscape (Ismagilova et al., 2022).

Another advantage lies in improved threat detection. Advanced algorithms have the capacity to analyze larger datasets, enabling the identification of subtle patterns that may indicate potential security threats. This heightened analytical capability has the potential to uncover novel attack methods, allowing for proactive security measures and the prevention of security incidents before they escalate.

The automation of data security tasks is facilitated by robust algorithms, leading to increased efficiency, reduced human error, and real-time threat response. This automation not only streamlines security operations but also ensures a more agile and adaptive defense against evolving cyber threats. Privacy-preserving analytics represent a significant advancement, allowing organizations to conduct data analysis while safeguarding sensitive information. Specialized algorithms enable the extraction of valuable insights without compromising individual privacy, striking a delicate balance between deriving meaningful information and protecting personal data (Cremer et al., 2022).

Furthermore, the development of new algorithms aligns with the imperative of compliance with data privacy regulations. As regulations evolve and become more stringent, organizations can leverage these advanced algorithms to ensure compliance and build trust with stakeholders. The adaptability of these algorithms to changing regulatory landscapes positions organizations to navigate the complex terrain of data privacy while maintaining a robust security framework. In summary, while the development of data security algorithms for Big Data introduces notable advantages, it is crucial to approach these advancements with a critical lens, acknowledging potential challenges and ensuring that security measures align with ethical considerations and regulatory requirements.

Despite the notable advantages of developing data security algorithms for Big Data, there are inherent disadvantages that warrant careful consideration. One critical concern is algorithmic bias, wherein security algorithms may exhibit biases that lead to discriminatory or unfair outcomes (Mvula et al., 2023). It is imperative to approach algorithm design and testing with diligence to mitigate this risk and ensure that security measures do not inadvertently perpetuate inequalities. Another significant drawback is the computational complexity associated with processing large datasets using sophisticated algorithms. This complexity often translates into computational expenses, requiring specialized hardware and expertise. The financial and technical investments necessary for managing this complexity may pose challenges for some organizations.

Limited transparency is a recurring issue with complex algorithms. The intricacies of these advanced security measures can be challenging to understand and explain, raising concerns about accountability and the potential for misuse. Striking a balance between robust security and the transparency needed for ethical and accountable practices becomes a delicate task. Privacy trade-offs are also a consideration, as while some algorithms anonymize data, others may still allow re-identification. The challenge lies in carefully balancing the imperatives of security and privacy to ensure that sensitive information remains protected even as analyses are conducted.

Furthermore, there is a shortage of a skilled workforce capable of implementing and managing advanced data security algorithms. The expertise required for the development, deployment, and maintenance of these algorithms can be scarce and expensive, posing a barrier to their widespread adoption (Sumithra & Parameswari, 2022). It is crucial to acknowledge that the specific advantages and disadvantages will vary depending on factors such as the type of data, industry, and the nature of security threats faced. Considering the contextual nuances is essential in tailoring data security strategies to the specific needs and challenges of an organization.

Continuous improvement is a fundamental principle in the realm of data security algorithms. These algorithms must undergo ongoing development and updates to stay abreast of evolving threats and technologies. Embracing this dynamic and adaptive approach is key to maintaining the effectiveness of security measures over time (Li et al., 2018). Finally, ethical considerations play a paramount role throughout the development and deployment process. Ensuring algorithmic fairness, transparency, and responsible use is essential for building and maintaining trust, both within organizations and among the broader community. In conclusion, by carefully weighing these advantages and disadvantages and incorporating important considerations, organizations can responsibly leverage data security algorithms to enhance information security in the era of Big Data.

CONCLUSION

In conclusion, the development of data security algorithms for Big Data represents a critical response to the escalating challenges posed by the unprecedented volume and complexity of contemporary datasets. The literature review provides a comprehensive exploration of the evolution of these algorithms, emphasizing the need for robust solutions tailored explicitly for the unique characteristics of Big Data. The ongoing efforts to address vulnerabilities, advance technologies, and incorporate cryptographic techniques underscore a commitment to establishing a resilient foundation for information security in the dynamic landscape of Big Data. As organizations grapple with the multifaceted challenges, the emergence of innovative data security algorithms serves as a beacon of progress, offering enhanced protection, improved threat detection, and privacy-preserving analytics.

However, the advantages of these advancements must be carefully balanced with the inherent disadvantages. Algorithmic bias, computational complexity, limited transparency, privacy trade-offs, and workforce shortages pose significant considerations. While the benefits include heightened data protection, improved threat detection, and privacy preservation, organizations must navigate challenges such as potential biases, resource-intensive processes, and the need for skilled professionals. Continuous improvement and ethical considerations are pivotal in ensuring responsible deployment and adaptation of these algorithms to the evolving landscape of Big Data. By addressing these complexities and embracing a comprehensive and adaptive approach, organizations can leverage data security algorithms responsibly to enhance information security in the era of massive and diverse datasets.

REFERENCES

Chandra, S., Ray, S., & Goswami, R. T. (2017). Big Data Security: Survey on Frameworks and Algorithms. *2017 IEEE 7th International Advance Computing Conference (IACC)*, 48–54. https://doi.org/10.1109/IACC.2017.0025

Cremer, F., Sheehan, B., Fortmann, M., Kia, A. N., Mullins, M., Murphy, F., & Materne, S. (2022). Cyber risk and cybersecurity: a systematic review of data availability. *The Geneva Papers on Risk and Insurance - Issues and Practice*, *47*(3), 698–736. https://doi.org/10.1057/s41288-022-00266-6

Dal Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S., & Bontempi, G. (2014). Learned lessons in credit card fraud detection from a practitioner perspective. *Expert Systems with Applications*, *41*(10), 4915–4928. https://doi.org/10.1016/j.eswa.2014.02.026

Georgiadis, G., & Poels, G. (2022). Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review. *Computer Law & Security Review*, *44*, 105640. https://doi.org/10.1016/j.clsr.2021.105640

Gheyas, I. A., & Abdallah, A. E. (2016). Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis. *Big Data Analytics*, *1*(1), 6. https://doi.org/10.1186/s41044-016-0006-0

Ismagilova, E., Hughes, L., Rana, N. P., & Dwivedi, Y. K. (2022). Security, Privacy and Risks Within Smart Cities: Literature Review and Development of a Smart City Interaction Framework. *Information Systems Frontiers*, *24*(2), 393–414. https://doi.org/10.1007/s10796-020-10044-1

Lei Xu, Chunxiao Jiang, Jian Wang, Jian Yuan, & Yong Ren. (2014). Information Security in Big Data: Privacy and Data Mining. *IEEE Access*, *2*, 1149–1176. https://doi.org/10.1109/ACCESS.2014.2362522

Li, X., Niu, J., Kumari, S., Wu, F., Sangaiah, A. K., & Choo, K.-K. R. (2018). A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments. *Journal of Network and Computer Applications*, *103*, 194–204. https://doi.org/10.1016/j.jnca.2017.07.001

Moghadam, R. S., & Colomo-Palacios, R. (2018). Information security governance in big data environments: A systematic mapping. *Procedia Computer Science*, *138*, 401–408. https://doi.org/10.1016/j.procs.2018.10.057

Mvula, P. K., Branco, P., Jourdan, G.-V., & Viktor, H. L. (2023). A systematic literature review of cyber-security data repositories and performance assessment metrics for semi-supervised learning. *Discover Data*, *1*(1), 4. https://doi.org/10.1007/s44248-023-00003-x

Rawat, D. B., Doku, R., & Garuba, M. (2021). Cybersecurity in Big Data Era: From Securing Big Data to Data-Driven Security. *IEEE Transactions on Services Computing*, *14*(6), 2055–2072. https://doi.org/10.1109/TSC.2019.2907247

Sharafaldin, I., Habibi Lashkari, A., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, 108–116. https://doi.org/10.5220/0006639801080116

Shaukat, K., Luo, S., Varadharajan, V., Hameed, I. A., & Xu, M. (2020). A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. *IEEE Access*, *8*, 222310–222354. https://doi.org/10.1109/ACCESS.2020.3041951

Singh, M., Halgamuge, M. N., Ekici, G., & Jayasekara, C. S. (2018). *A Review on Security and Privacy Challenges of Big Data* (pp. 175–200). https://doi.org/10.1007/978-3-319-70688-7_8

Sumithra, R., & Parameswari, R. (2022). Data privacy and data protection security algorithms for big data in cloud. *International Journal of Health Sciences*, 7613–7621. https://doi.org/10.53730/ijhs.v6nS2.6834