IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 2, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, April 2024 ||

International Conference on Recent Development in Engineering and Technology – ICRDET 24

Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, Tamilnadu, India

Security and Privacy for Big Data

Mrs. P. Anitha, M.E, S. Shabarie, R. Devadharshini, S. Barathkumar

Assistant Professor, Department of InformationTechnology, Paavai Engineering College (Autonomous), Namakkal,Tamil Nadu, India

Department of Information Technology, Paavai Engineering College (Autonomous), Namakkal, Tamil Nadu, India

Department of Information Technology,Paavai Engineering College (Autonomous), Namakkal, Tamil Nadu, India

Department of Information Technology,Paavai Engineering College (Autonomous), Namakkal, Tamil Nadu, India

ABSTRACT: Big data has changed the way organizations collect, process and analyze big data.However, in addition to their benefits, they also bring important security and privacy issues. This article examines the various security and privacy issues inherent in big data and explores solutions to mitigate these risks. By solving these problems, organizations can leverage the power of big data and at the same time secure the confidentiality, integrity and availability of sensitive data.

I. INTRODUCTION

In today's business world, big data has become a transformative force that is changing the way organizations collect, process and analyze big data to obtain useful information. Big data reveals hidden patterns, trends, and relationships, allowingcompanies to make decisions, improve processes, and get better results in today's business world. However, the proliferation of big data, in addition to the numerous benefits it provides, also brings security and privacy issues that need to be carefully evaluated. Security and privacy issues are particularly important in the field of big data,mainly due to the volume and sensitivity of the data involved. Large data databases thatoften contain personally identifiable information (PII), financial information, andbusiness information provide good targets for criminals who cannot gain access. The possible consequences of data breaches and unauthorized access range from financial loss and reputational damage to liability and damage to customer trust.

The distribution and interaction of big data creates complex security and privacyissues. The information spans multiple platforms, networks, and repositories, leading to potential disruption and exploitation. The evolving threat landscape, characterized by advanced cyber attacks and insider threats, highlights the need for security measures and careful monitoring in the big data ecosystem. According to these challenges, organizations need to ensure security to secure big data. Integrity, confidentiality and availability of information in big data. Companies can strengthen their defenses against external threats and internal risks by implementing effective security measures, encryption methods, access control and privacy-enhancing technologies. Additionally, compliance with regulatory requirements such as GDPR, CCPA and HIPAA ensures compliance with data protection laws and supports a culture of transparency and accountability.

As big data continues to shape the path of business innovation and digital transformation, addressing security and privacy issues remains critical. By implementing the best approach to data governance and risk management, organizations can take full advantage of big data while maintaining stakeholder trust. Pressure and trust in the information world is increasing.

II. SECURITY CHALLENGES IN BIG DATA

Securing big data is a huge challenge and risks are everywhere. The large amountof sensitive data stored in these large systems makes them a prime target for cyber attacks. Privacy laws such as GDPR and CCPA are complex and require strong access protections and controls to protect personal data. Regulatory challenges are mounting, with disparate sources and structures creating a complex storm that requires clear policies to address. As data volume increases like a rising tide, scalability becomes important to prevent overflow without compromising security. Internal threats lurk in the shadows, attack at lightning speed, and require tight control and careful monitoring.Data sovereignty issues transcend national borders and must be pursued by internationallaw. Cybersecurity threats create a storm of malware and phishing

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, April 2024 ||

International Conference on Recent Development in Engineering and Technology – ICRDET 24

Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, Tamilnadu, India

attacks; Therefore, you need to strengthen your defenses and stay alert. A storm of complexity and integration is coming, it is difficult to integrate and needs to be carefully coordinated. Finally, protecting information throughout its lifecycle is a complex task that requires ongoing protection against access and control at all levels. Solving these challenges requires a strategic approach that combines efficiency, strong policies, and constant vigilance against storms and protecting data. , shown in figure 2.

Fig : 2 Security Challenges In Big Data

2.1 Information leakage and unauthorized access

Information leakage and unauthorized access pose a serious risk to big data. Unauthorized access to sensitive information can lead to a variety of consequences, including financial loss, reputational damage, liability, and customer loss. Hackers, malicious actors, and inadequate security measures can contribute to this vulnerability.Organizations need to implement effective authentication, authorization, and encryption measures to mitigate these risks.

2.2 Data Integrity

Ensuring data integrity throughout the life cycle is very important in maintainingthe reliability and integrity of the data. Data may be subject to corruption, tampering, or unauthorized alteration, which could compromise its accuracy and reliability. Organizations should use data validation, checksum, and error correction procedures todetect and prevent data problems. Additionally, backup, redundancy, and data validation procedures must be in place to ensure data consistency and reliability.

2.3 Insider Threats

Malicious insiders pose a serious risk to big data because they have access to big data. Important information and systems. Insider threats may include employees, contractors, or business partners who maliciously or unintentionally compromiseinformation security. Organizations must implement robust controls, employee monitoring, and security training to protect and defend against insider threats. Additionally, using minimum rules and separation of duties can help reduce the risk ofinsider attacks.

2.4 Distribution of big data

Distribution of big data brings with it problems regarding data protection in different areas. platform and environment. Data fragmentation, duplication, and synchronization of distributed systems increases the risk of inaccessibility and data leakage. Organizations need to implement strong encryption, authentication, and management systems to protect

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, April 2024 ||

International Conference on Recent Development in Engineering and Technology – ICRDET 24

Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, Tamilnadu, India

data in transit and at rest. Additionally, network segmentation, firewalls, and intrusion detection tools can help reduce security risks onlarge data sets.

III. PRIVACY ISSUES IN BIG DATA

3.1 Protection of Personally Identifiable Information (PII)

Protection of Personally Identifiable Information (PII) is important to comply with and maintain data privacy laws Consumer Trust is very important. Big data often collects and processes large amounts of personal information; This makes them attractive targets for hackers and data breaches. Organizations must use strong encryption, access control, and data anonymization technologies to effectively protect PII. Additionally, compliance with data privacy laws such as GDPR and CCPA requires organizations to obtain explicit consent to collect and process personal information and provide transparency in the use of such data.

3.2 De-identification and Anonymization

De-identification and Anonymization are important to protect privacy when managing the electronic use of information for analysis and research purposes. This process removes or highlights suspicious information from the dataset, thus reducing the risk of re-identification while preserving the value of the data. Organizations shoulduse strategies such as k-anonymity, l-diversity, and differential privacy to effectively anonymize data. Additionally, using data depersonalization, tokenization, andpseudonymization techniques can improve privacy protection in big data.

3.3 Privacy-preserving data mining

Privacy-preserving data mining technology allows organizations to analyze sensitive data without compromising personal privacy. This technology uses encryption and privacy-enhancing technology to process information while preserving privacy.The security of multiparty computing, homomorphic encryption, and state learning areexamples of data privacy protection. Organizations need to use this technology to comply with data privacy laws and protect individual privacy rights.

3.4 User Consent and Consent

It is essential to obtain consent from users and ensure transparency regarding thecollection and use of information. Fear of complying with personal information. Organizations must obtain explicit consent from users before collecting and processingtheir personal data and provide a privacy policy regarding the use of the data. Additionally, organizations must be transparent about their data collection, storage, anduse practices to build trust with users and be accountable for protecting their privacy.

IV. SOLUTIONS AND BEST PRACTICES

4.1 Encryption

Encryption is a simple security measure to protect data at rest and in transit. Organizations need to implement strong encryption algorithms and procedures toencrypt sensitive data and prevent unauthorized access. Additionally, the use of encryption key management practices and secure encryption protocols can increase datasecurity in large databases.

4.2 Access control mechanisms

Access control mechanisms such as role-based access control (RBAC) and attribute-based access control (ABAC) are sensitive to constraints in big data systems. important information. Organizations must follow access control policies and procedures to maintain minimum access rights and separate activities. Additionally, using user authentication, authorization, and auditing systems can help enforce access control policies and detect unauthorized access attempts.

4.3 Data Masking and Tokenization

Data masking and tokenization technologies can protect sensitive data while protecting valuable data. Organizations should use data masking techniques to anonymize sensitive data by replacing original values with masked values. Additionally, using encryption technology to replace sensitive data with generated identifiers can further enhance data security. Organizations should use data depersonalization and tokenization techniques to protect sensitive data in big data.

4.4 Privacy-Enhancing Technologies

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, April 2024 ||

International Conference on Recent Development in Engineering and Technology – ICRDET 24

Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, Tamilnadu, India

Privacy-enhancing technologies such as differential privacy and homomorphic encryption are new technologies that enable organizations to protect privacy in big dataanalytics. While different privacy techniques introduce noise into the query to protect sensitive data, homomorphic encryption can perform calculations on encrypted data without decrypting the data. Organizations must implement privacy-enhancingtechnology to comply with data privacy laws and protect individual privacy rights.

V. CASE STUDIES

5.1 Real World Examples of Security and Privacy Issues

Several Real World Examples report on security and privacy issues in big data. Examples include data breaches, unauthorized access, insider threats, and privacy breaches. Organizations should learn from these examples and implement best practicesto reduce security and personal risk.

5.2 Lessons Learned and Best Practices

Analysis of real-life situations can provide insight into lessons learned and practices: best practices for solving security and privacy issues in big data. Organizations need to implement security protection, privacy-enhancing technology, and compliance to protect sensitive information and protect people's privacy.

Security and privacy issues in big data require a comprehensive approach that includes the design of security measures, implementation of privacy-enhancing technology and compliance procedures. By implementing solutions and best practices, organizations can reduce security and privacy risks and build trust among users and stakeholders, respectively. Additionally, analysis of real-world trends can provide important lessons and best practices for resolving security and privacy issues in big data**.**

VI. FUTURE DIRECTIONS

6.1 Professional Development Discussion More Information More :

The rise of artificial intelligence (AI) and machine learning (ML) to improve threat intelligence and predictive analytics is one example. AI-driven security solutions can now analyze big data to identify unusual behavioral patterns and potential vulnerabilities, thus improving protection protection in big data. Another emerging trend is the growth of computing and Internet of Things (IoT) devices; This brings newchallenges and opportunities to protect data at the edge of the network. Edge computingreduces latency and increases efficiency by allowing data processing and analysis to occur closer to the data source. However, protecting the data created and transmitted by IoT devices presents unique challenges due to its nature and limited resources. Future advances in security technology, such as blockchain-based authentication and lightweight encryption protocols, will be important to maintain data integrity and confidentiality in the business environment.

Advances in encryption technology such as homomorphic encryption and the security of multi-party computing will improve data privacy in big data analytics.Homomorphic encryption allows direct computation of encrypted data without decryption, allowing data analysis while preserving privacy. Similarly, the security of multiparty computing allows many people to come together to compute the work of their ideas while preserving their privacy, thus opening new possibilities for information collaboration without compromising privacy. Adoption of privacy tools such as privacy diversity and government education will continue to provide benefits as organizations seek to balance data use and privacy protection. Different privacy systems add noise tothe question to protect personal privacy while still providing accurate information. Federated learning allows models to be trained on separate datasets without sharing raw data, increasing model accuracy while protecting data privacy.

6.3 Identify areas for further research and development

Despite advances in big data and privacy, there are still unaddressed issues that require further research and development. Work. One of the areas of focus is the development of strong authentication and management systems appropriate to the nature of big data. Traditional access control models can struggle to adapt to the scale and complexity of big data, requiring new methods such as attribute-based access control (ABAC) and risk management.

Another area of research is the development of secure information sharing systems that facilitate collaboration while protecting information privacy and confidentiality. Securing data sharing is crucial for organizations to derive insights

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, April 2024 ||

International Conference on Recent Development in Engineering and Technology – ICRDET 24

Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, Tamilnadu, India

fromdisparate data while maintaining data ownership and privacy concerns. Future researchshould explore strategies for secure data exchange, data anonymization, and trust discussions to enable shared and secure data across multiple environments. Advances in data anonymization and identification technologies are needed to meet changing privacy and regulatory requirements. Current anonymization techniques may not be sufficient to resist re-identification attacks or store useful data, so stronger anonymization algorithms and privacy-improved technologies must be developed.

VII. CONCLUSION

Big data faces significant security and privacy issues that require organizations to think carefully and take protection measures. The large amounts of data collected andprocessed on these systems make them a valuable target for criminals seeking access, increasing the likelihood of crimes and private crimes. Additionally, the distribution of big data environments and the connectivity of data across multiple platforms increases the complexity of security threats and requires security measures designed to mitigate risks. Ensuring the security and confidentiality of big data requires good security procedures, encryption procedures, access control and security. Organizations must adopt a multi-layered security approach that includes understanding sensitive data, stringent controls under the law of least privilege, and continuous monitoring of unauthorized activity. Additionally, using privacy-enhancing techniques such as differential privacy and homomorphic encryption can help protect private information while providing meaningful insights from file size.

Organizations must comply with data protection laws such as GDPR, CCPA, and HIPAA to ensure that sensitive information is processed lawfully and legally. This includes obtaining explicit consent from users to collect and process data, ensuring transparency about the use of data, and using anonymized data to protect privacy. Privacy Policy. Based on these challenges and recommendations, organizations should invest in building security and privacy awareness, providing training to employees andstakeholders, and encouraging collaboration between IT, security, and compliance teams. By implementing security measures, privacy-enhancing technologies, and compliance procedures, organizations can improve the security and privacy of their datacapital, protect sensitive information, and maintain stakeholders' trust and compliance with information in a world of trust.

REFERENCES

[1] Ninny Bhogal, Shaveta Jain," A Review on Big Data Security and Handling", International Research Based Journal, Vol(6)-Issue(1) , ISSN 2348-1943, March,11, 2017.
[2] Mohammed S.Al-Kahtani," Security and Privacy in Big Data", International Journalof Computer Engineering and Information Technology, VOL. 9, NO. 2, E-ISSN 2412-8856, February 2017.
[3] Prof. Amar Nath Singh, Er. Anurag Pattanayak, Er. Gyanachanda Samantaray," Data Analytics Application used in the field of Big Data for Security Intelligence", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 6, Issue 1, ISSN 2278- 6856, January - February 2017.
[4] Minit Arora, Dr Himanshu Bahuguna, "Big Data Security – The Big Challenge", International Journal of Scientific & Engineering Research, Volume 7, Issue 12, ISSN 2229-5518, December-2016.
[5] Naveen Rishishwar, Vartika, Mr. Kapil Tomar," Big Data: Security Issues and Challenges", International Journal of Technical Research and Applications, e-ISSN: 2320-8163, Special Issue 42 (AMBALIKA), PP. 21-25, March 2017.
[6] Bhavani Thuraisingham, "Big Data – Security with Privacy", NSF Workshop, September 16-17, 2014.
[7] Trupti V. Pathrabe, "Survey on Security Issues of Growing Technology: Big Data",IJIRST, National Conference on Latest Trends in Networking and Cyber Security, March 2017.
[8] Venkata Narasimha Inukollu, Sailaja Arsi and Srinivasa Rao Ravuri," Security issues associated with Big Data in Cloud Computing", International Journal of NetworkSecurity & Its Applications (IJNSA), Vol.6, No.3, May 2014.
[9] R.Kalaivani," Security Perspectives on Deployment of Big Data using Cloud: A Survey", International Journal of Advanced Networking & Applications (IJANA), Volume: 08, Issue: 05 Pages: 5-9, Special Issue, 2017.
[10] K.P.Maheswari, P.Ramya, S.Nirmala Devi," Study and Analyses of Security Levels in Big Data and Cloud Computing", International on Recent Trends in Engineering Science, Humanities and Management, February 2017.
[11] Mr. Shrikant Rangrao Kadam, Vijaykumar Patil, "Review on Big Data Security in Hadoop", International Research Journal of Engineering and Technology (IRJET), eISSN: 2395 -0056, Volume: 04 Issue: 01, pISSN: 2395-0072, Jan -2017.
[12] Philip Derbeko, Shlomi Dolev, Ehud Gudes, Shantanu Sharma, "Security and Privacy Aspects in MapReduce

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, April 2024 ||

International Conference on Recent Development in Engineering and Technology – ICRDET 24

Organized by

Dhaanish Ahmed Institute of Technology, KG Chavadi, Coimbatore, Tamilnadu, India

on Clouds: A Survey", Elsevier Computer Science Review, arXiv:1605.00677v1 [cs.DB] 2 May 2016.
[13] Asha Patel, "A Survey Paper on Security Issue with Big Data on Association Rule Mining", IJIRST, National Conference on Latest Trends in Networking and Cyber Security, March 2017.
[14] Vinod B. Bharat, Pramod B. Deshmukh, Laxmikant S. Malphedwar, P. Malathi and Nilesh N. Wani, "Big Data and Database Security", IJCTA, 10(8), pp. 517-528 ISSN: 0974-5572, International Science Press, 2017.
[15] Sanchita Gupta, Akashkataria, Shubham Rathore, Dharmendra Singh Rajput," Information Security Issues in Big Data: Solution using PPDM (Privacy Preserving Data Mining)", International Journal of Pharmacy & Technology, ISSN: 0975-766X, Vol. 8, Issue No.4, November 2016.
[16] Thayananthan V, Albeshri A. Big data security issues based on quantum cryptography and privacy with authentication for mobile data center. Procedia Computer Science. 2015 Jan 1; 50: 149- 56.
[17] Yu S. Big privacy: Challenges and opportunities of privacy study in the age of big data. IEEE access. 2016; 4: 2751-63.
[18] Demchenko Y, Ngo C, de Laat C, Membrey P, Gordijenko D. Big security for big data: Addressing security challenges for the big data infrastructure. In Workshop on Secure Data Management 2013 Aug 30 (pp. 76-94). Springer, Cham.
[19] Hu J, Vasilakos AV. Energy big data analytics and security: challenges and opportunities. IEEE Transactions on Smart Grid. 2016 Sep;7(5):2423- 36.
[20] Mehmood A, Natgunanathan I, Xiang Y, Hua G, Guo S. Protection of big data privacy. IEEE access. 2016; 4:1821-34.
[21] Gadepally V, Hancock B, Kaiser B, Kepner J, Michaleas P, Varia M, Yerukhimovich A. Computing on masked data to improve the security of big data. In2015 IEEE International Symposium on Technologies for Homeland Security (HST)2015 Apr 14 (pp. 1-6). IEEE.
[22] Sharma PP, Navdeti CP. Securing big data hadoop: a review of security issues, threats and solution. Int. J. Comput. Sci. Inf. Technol. 2014; 5 (2):2126-31.
[23] Richards NM, King JH. Three paradoxes of big data. Stan. L. Rev. Online. 2013; 66: 41.
[24] Bharati, T. S. (2015). Enhanced Intrusion Detection System for Mobile Adhoc Networks using Mobile Agents with no Manager. International Journal of Computer Applications, 111(10).
[25] Bharati, T. S., & Kumar, R. (2015, March). Secure intrusion detection system for mobile adhoc networks. In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on (pp. 1257-1261). IEEE.
[26] Bharati, T. S., & Kumar, R. (2015). Intrusion Detection System for MANET using Machine Learning and State Transition Analysis. International Journal of Computer Engineering & Technology (IJCET), 6(12), 1-8.
[27] Bharati, T. S., & Kumar, R. (2016). Enhanced Key Distribution for Mobile Adhoc Networks. International Journal of Engineering Science, 6(4), 4184-4187.
[28] Bharati T. S. (2017). Agents to Secure MANETS. International Journal of Advanced Engineering and Research Development, 4(11), 1267-1273.
[29] Bharati T.S. (2018). MANETs and Its' Security. International Journal of Computer Networks and Wireless Communication, 8(4), 166-171.
[30] Jaseena KU, David JM. Issues, challenges, and solutions: big data mining. CS & IT-CSCP. 2014 Dec 27; 4 (13):131-40.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details