SOFTWARE DESIGN & DEVELOPMENT

Computer Network & Security

Title: Securing Big Data: Network Attacks and Access Control Challenges

**1.0 Introduction to Big Data**

According to the research (Batko and Ślęzak, 2022), Big data refers to massive volumes of structured and unstructured data that organizations collect and process on a daily basis to gain insights and make informed decisions. This data often comes from various sources, including business applications, web logs, social media, sensors, transactions, and more (Batko and Ślęzak, 2022). The importance of big data lies in its potential to drive innovation, improve business operations, and enhance decision-making processes. However, the immense volume, velocity, and variety of big data also pose significant security challenges, making it an attractive target for cyberattacks (Hossain et al., 2019).

**1.1.0 Big Data Attacks**

Research by (Suraj, Kumar Singh and Tomar, 2018) argues that Big data systems are vulnerable to a range of cyberattacks due to their distributed nature and the immense amount of valuable information they store. Common big data attacks include:

* Data Theft: Attackers may attempt to steal sensitive data stored within big data systems, such as customer information, financial records, or intellectual property. This can lead to severe financial losses and reputational damage.
* Data Breaches: Adversaries often target big data repositories to steal sensitive information, such as customer data, financial records, and intellectual property. Data breaches can result in severe financial losses and damage to an organization's reputation.
* Data Manipulation: Adversaries might alter or manipulate data within big data repositories, leading to incorrect business decisions, loss of trust in data integrity, and operational disruptions.
* Denial of Service (DoS) Attacks: These attacks aim to overload big data systems, rendering them unavailable to legitimate users. As a result, critical business processes may come to a halt.
* Insider Threats: Malicious or negligent insiders can misuse their access privileges to compromise the confidentiality, integrity, or availability of big data resources.

**1.1.1 Threat Analysis**

Threat analysis is a crucial step in safeguarding big data systems. It involves identifying potential threats, assessing their impact and likelihood, and developing strategies to mitigate them(Georgiadis and Poels, 2022). In the context of big data security:

* Threat Identification: Organizations must continually monitor and analyze their big data environments to identify emerging threats. This includes scrutinizing access logs, network traffic, and system behavior for anomalies.
* Risk Assessment: After identifying threats, it's essential to assess their potential impact on the organization. This involves evaluating the financial, operational, and reputational consequences of a successful attack.
* Mitigation Strategies: Organizations should implement robust access controls, encryption, and intrusion detection systems to reduce the risk of big data breaches. Regular security audits and penetration testing can help identify and address vulnerabilities.
* Incident Response: Establishing an incident response plan is crucial for quickly detecting and mitigating big data breaches when they occur. This plan should outline the steps to take when an incident is detected, including containment, investigation, and recovery procedures.

Securing big data against network attacks is a complex yet essential task for organizations. Understanding the nature of big data, recognizing potential threats, and implementing effective access controls and threat analysis measures are critical steps towards safeguarding this valuable asset and mitigating the associated risks.

**2.0 Unique/New Access Control Challenges for Securing Big Data**

Research by (Hossain et al., 2019) argues that the advent of big data has revolutionized the way organizations store, manage, and utilize data. In an era characterized by the relentless generation and accumulation of vast datasets, the security of big data has emerged as a paramount concern for organizations worldwide. Big data repositories often contain sensitive information, making them attractive targets for cyber adversaries (Rawat, Doku and Garuba, 2019). While access control is a fundamental component of data security, the unique characteristics of big data introduce new challenges that demand innovative solutions. This section delves into the distinct access control challenges inherent in securing big data and proposes novel solutions to mitigate risks and enhance data security.

**2.1.0 Data Volume and Velocity**

Big data systems generate and process vast amounts of data at unprecedented speeds (Bakshi, 2012). Traditional access control mechanisms struggle to keep pace with the volume and velocity of data. The sheer quantity of data makes it challenging to define and manage access permissions effectively. Moreover, real-time analysis and access control are essential to detect and respond to threats promptly (Bakshi, 2012).

Bakshi (2012) argues that implementing scalable access control solutions that can handle high data volumes and provide real-time access decisions is crucial. Technologies like distributed access control lists (ACLs) and in-memory databases can be employed to enhance access control performance.

**2.1.1 Data Variety**

In line with the previous research (Lo Giudice et al., 2019), Big data encompasses a wide range of data types, including structured, semi-structured, and unstructured data. This diversity poses a challenge for traditional access control models that are primarily designed for structured data sources.

Implementing data classification and tagging is essential to differentiate and apply access controls effectively. Attribute-based access control (ABAC) and policy-based access control can be used to manage access to different data types based on their attributes (Gupta, Patwa and Sandhu, 2018).

**2.1.2 Data Storage and Replication**

Big data often resides in distributed storage systems and undergoes replication for fault tolerance and high availability. Ensuring consistent access control across multiple copies of data is a complex task. Inconsistent access controls can lead to data leakage and unauthorized access.

Employ distributed access control mechanisms that synchronize access policies across data replicas. Blockchain-based access control systems can provide a tamper-resistant mechanism for ensuring consistency (Rana et al., 2022).

**2.1.3 Scalability**

Big data environments are highly scalable, with the ability to add and remove nodes dynamically. Traditional access control solutions struggle to scale seamlessly with the environment, leading to management overhead and potential access control gaps.

Implementation of elastic access control solutions that can automatically scale with the environment. Containerized access control systems and microservices architecture can facilitate scalability (Casalicchio and Iannucci, 2020).

**2.1.4 User and Data Attribution**

Identifying users and their associated data in a big data ecosystem can be challenging. Users and data may be spread across multiple platforms, making it difficult to establish a clear link between them.

Implement identity and data mapping techniques to establish user-data relationships. Utilize identity federation and single sign-on (SSO) solutions to enhance user attribution (Cao, 2023).

**2.1.5 Privacy Concerns**

Big data often contains sensitive and personally identifiable information (PII). Ensuring privacy while providing access controls is a delicate balance (Koo, Kang and Kim, 2020). Traditional access control models may not adequately address privacy concerns.

Implement fine-grained access control and encryption to protect sensitive data. Differential privacy techniques can be used to anonymize data while allowing limited access(Koo, Kang and Kim, 2020).

**2.2.0 Innovative Approaches to Access Control for Big Data**

**2.2.1 Attribute-Based Access Control (ABAC)**

As stated by Hu et al. (2014), ABAC is a dynamic access control model that uses attributes to make access decisions. The attributes considered by ABAC to access control includes users, data, and the environment. It is well-suited for managing diverse data types in big data environments. It allows for fine-grained control over data access based on multiple attributes, addressing the variety and velocity challenges. ABAC policies can consider user attributes, data attributes, and environmental attributes to grant or deny access.

**2.2.2 Role-Based Access Control (RBAC) with Dynamic Roles**

RBAC can be enhanced by introducing dynamic roles that adapt to changing user roles and data access needs. Dynamic roles can be based on user behavior and data usage patterns, ensuring that access remains relevant and controlled (Agrawal et al., 2011).

**2.2.3 Data Encryption**

Implementing encryption at rest and in transit can provide an additional layer of security for big data. Encrypted data is less susceptible to unauthorized access even if perimeter defenses are breached (Agrawal et al., 2011).

**2.2.4 Machine Learning and AI-Powered Access Control**

Leveraging machine learning and artificial intelligence can enhance access control in big data environments. These technologies can analyze user behavior patterns, detect anomalies, and trigger alerts or access restrictions in real time. This enables proactive threat detection and adaptive access control decisions based on user behavior.

**2.2.5 Data Masking and Redaction**

To protect sensitive information while maintaining data usability, data masking and redaction techniques can be employed (Agrawal et al., 2011). These methods replace sensitive data with fictional or pseudonymous values for certain users or applications.

**2.2.6 Cloud Security Solutions**

According to Balachandran and Prasad (2017), many big data deployments are on cloud platforms. Cloud security solutions offer robust access control features tailored to cloud-based environments, including identity and access management (IAM) services.

**2.2.7 Blockchain for Access Control**

According to Rana et al. (2022), Blockchain technology can be used to enhance access control by providing a decentralized, immutable ledger of access permissions. Blockchain technology can also be used to create tamper-resistant access control policies and audit trails. Access control decisions can be recorded on a blockchain. This can increase transparency, immutability, and trust in access control decisions (Rana et al., 2022).

Research by (OUTCHAKOUCHT, ES-SAMAALI and Philippe, 2017) suggests that, securing big data against unauthorized access is a multifaceted challenge due to its volume, variety, velocity, and complexity. Traditional access control mechanisms are often ill-equipped to handle these challenges. Innovative approaches, such as ABAC, encryption, machine learning, and blockchain, offer promising solutions to address the unique access control challenges associated with big data (OUTCHAKOUCHT, ES-SAMAALI and Philippe, 2017). Organizations must adopt a holistic approach that combines these techniques to safeguard their valuable data assets and mitigate the risks posed by cyber adversaries.

**3.0 Critically discussion of existing approaches for securing big data**

The era of big data has ushered in a wealth of opportunities for organizations to leverage data-driven insights. However, the growing volume, variety, and velocity of data have also raised significant security concerns. Securing big data is crucial to prevent unauthorized access, data breaches, and potential financial losses. This section critically discusses existing approaches for securing big data, evaluating their strengths and weaknesses.

**3.1 Traditional Access Control Methods**

**3.1.0 Role-Based Access Control (RBAC)**

**Strengths**

RBAC is a well-established access control model that has been widely adopted across various industries and organizations (Aftab et al., 2015). One of its key strengths lies in its familiarity and proven track record, making it a trusted choice for access management.

Additionally, RBAC simplifies the process of access management by associating permissions with roles. This approach streamlines the assignment of access rights, reducing the complexity of managing permissions for individual users. By defining access based on job functions or roles, RBAC also offers a level of granularity that aligns well with organizational hierarchies and responsibilities(Aftab et al., 2015).

Furthermore, RBAC provides a structured framework for access control, which is particularly beneficial in organizations with well-defined job functions and access requirements. This structured approach makes it easier to enforce security policies and maintain consistency in access management (Aftab et al., 2015).

**Weaknesses**

Despite its strengths, RBAC has limitations. One significant weakness is its limited capability to handle complex data types and fine-grained access control. RBAC primarily focuses on defining access based on roles, which may not adequately address the intricacies of data access requirements in scenarios where more granular control is necessary(Aftab et al., 2015).

Another weakness of RBAC is its lack of adaptability to dynamic changes in user roles and data access requirements. In dynamic organizations or projects where roles frequently change, adapting RBAC can become cumbersome, leading to potential inefficiencies and delays in access provisioning (Aftab et al., 2015).

Moreover, RBAC can lead to role explosion when managing access for large datasets or diverse user groups. As organizations grow and data becomes more diverse, the proliferation of roles to accommodate specific access needs can become overwhelming, making RBAC harder to manage effectively (Aftab et al., 2015).

**3.1.1 Discretionary Access Control (DAC)**

**Strengths**

Discretionary Access Control (DAC) is a flexible access control model that empowers data owners with the authority to control access to their data (Dattatray, Assistant and Sulabha, 2014). One of its key strengths lies in this decentralized approach to access management, which allows individuals or data owners to determine who can access their resources. This feature is particularly valuable in scenarios where data owners have unique knowledge about the sensitivity and access requirements of their data(Dattatray, Assistant and Sulabha, 2014).

Furthermore, DAC offers flexibility in defining access rules. It allows data owners to specify access permissions based on their discretion, accommodating a wide range of access scenarios. This flexibility can be advantageous in environments where specific access requirements vary widely among different data resources(Dattatray, Assistant and Sulabha, 2014).

**Weaknesses**

Despite its strengths, DAC has vulnerabilities that need to be considered. One significant weakness is its susceptibility to misconfigurations and unintentional data exposure(Dattatray, Assistant and Sulabha, 2014). Since DAC relies on data owners to make access decisions, errors in access permission assignments or misunderstandings about data sensitivity can lead to data breaches or unauthorized access.

DAC may not scale efficiently for big data systems with numerous users and data sources. In large and complex data environments, managing access permissions individually for each data resource can become unwieldy and time-consuming. The decentralized nature of DAC can hinder centralized control and oversight(Dattatray, Assistant and Sulabha, 2014).

Moreover, enforcing uniform security policies across diverse data types can be challenging with DAC. In environments with a wide variety of data formats and access requirements, ensuring consistent and comprehensive security policies may require significant effort. This can lead to gaps in security coverage and inconsistencies in access control(Dattatray, Assistant and Sulabha, 2014).

**3.2 Innovative Approaches**

**3.2.0 Attribute-Based Access Control (ABAC)**

**Strengths**

Attribute-Based Access Control (ABAC) offers a sophisticated and fine-grained approach to access control, leveraging user and data attributes for access decision-making (Hu et al., n.d.). One of its key strengths is its ability to provide granular access control based on a wide range of attributes, including user characteristics, data properties, and environmental factors. This granularity allows organizations to precisely tailor access permissions to meet specific requirements, making it well-suited for managing access in big data environments with diverse data types and complex access needs(Hu et al., n.d.).

ABAC's adaptability is another notable strength. It supports dynamic policy adaptation based on changing attributes. This means that access policies can evolve as attributes change over time, allowing organizations to maintain a high level of security and flexibility(Hu et al., n.d.). For example, access to data can be modified based on evolving user roles or changing data classifications, ensuring that access control remains aligned with the dynamic nature of big data systems(Hu et al., n.d.).

**Weaknesses**

ABAC's strengths are accompanied by certain weaknesses that need to be considered. First, implementing ABAC effectively requires a robust attribute management system. Managing and maintaining a comprehensive set of attributes for users and data resources can be complex and resource-intensive(Hu et al., n.d.). Organizations must establish clear processes for attribute management to ensure accuracy and consistency.

The complexity of policy definition and enforcement is another challenge associated with ABAC. Creating and managing access policies based on a wide range of attributes can be intricate, potentially leading to policy conflicts or ambiguities(Hu et al., n.d.). Furthermore, the enforcement of these policies can require a sophisticated access control infrastructure, which may pose challenges for organizations with limited resources or legacy systems.

**3.2.1 Data Masking and Anonymization**

**Strengths**

Data masking and anonymization are effective techniques for protecting sensitive data within big data environments(Marques and Bernardino, 2020). One of their primary strengths is their ability to safeguard sensitive information by replacing it with masked or anonymized values. This approach ensures that even if unauthorized access occurs, the actual sensitive data remains concealed, reducing the risk of data breaches and privacy violations.

These techniques excel in preserving data privacy while allowing controlled access. By masking or anonymizing data, organizations can strike a balance between data security and usability. Authorized users can still work with the data for legitimate purposes, such as analytics or testing, without exposure to sensitive details(Marques and Bernardino, 2020).

**Weaknesses**

Despite their strengths, data masking and anonymization have certain limitations that should be considered. One significant weakness is the potential for data loss and reduced analytical value. Masking or anonymizing data can obscure meaningful patterns and relationships within the dataset, limiting the insights that can be gained from analytics and data mining(Marques and Bernardino, 2020). Striking the right balance between data security and analytical utility can be challenging.

Data de-anonymization risks exist if these techniques are not implemented correctly. Skilled attackers may reverse-engineer anonymized data to reveal sensitive information, especially when additional data sources are available for correlation(Marques and Bernardino, 2020). This underscores the importance of careful implementation and adherence to best practices in data masking and anonymization.

**3.2.2 Encryption and Tokenization**

**Strengths**

Encryption and tokenization are robust data protection techniques with several significant strengths. First and foremost, they provide strong protection for data, both at rest and in transit. When data is encrypted or tokenized, it becomes unreadable to unauthorized individuals, ensuring that even if a breach occurs, the data remains confidential and secure(Obaidat et al., 2020b).

Another key strength is that encryption and tokenization ensure the confidentiality and integrity of data. Encryption algorithms use mathematical processes to scramble data into ciphertext, which can only be decrypted with the appropriate decryption key. This guarantees that sensitive information remains private and unaltered during transmission or storage.

**Weaknesses**

While encryption and tokenization offer robust data protection, they also come with certain weaknesses. One of the primary challenges is key management complexities. The security of encryption and tokenization heavily relies on managing encryption keys and tokens effectively. This involves generating, storing, and securely distributing keys, which can be a complex and resource-intensive process, especially in large-scale environments(Obaidat et al., 2020b).

Another potential drawback is the potential for performance overhead, particularly in real-time processing scenarios. The computational processes involved in encrypting and decrypting data or replacing it with tokens can introduce latency and impact system performance(Obaidat et al., 2020b). Organizations must carefully assess the trade-off between security and performance to ensure that the chosen approach aligns with their operational needs.

**3.3 Emerging Technologies**

**3.3.0 Homomorphic Encryption**

**Strengths**

Homomorphic encryption is an innovative cryptographic technique with several notable strengths. One of its key advantages is its ability to allow computation on encrypted data without requiring decryption. This unique feature enables organizations to perform data analysis and operations on sensitive data while it remains encrypted, ensuring that privacy is maintained throughout the entire process.

Another significant strength of homomorphic encryption is its provision of strong data privacy. By preserving data confidentiality during computation, homomorphic encryption offers robust protection against unauthorized access and data breaches. This is particularly valuable in scenarios where sensitive data must be outsourced for processing or analysis.

**Weaknesses**

Despite its strengths, homomorphic encryption comes with several limitations and challenges. One significant drawback is the high computational overhead associated with performing operations on encrypted data. The encryption and decryption processes, as well as the computation itself, can be resource-intensive and time-consuming. This can impact system performance, especially in real-time or high-throughput environments.

Limited practical implementations and adoption are also concerns for homomorphic encryption. While the concept is promising, practical implementations have been complex and not as widely adopted as other encryption methods. The specialized expertise required to deploy and manage homomorphic encryption systems can be a barrier to adoption for many organizations.

**4.0 Conclusion**

In conclusion, securing big data is of paramount importance due to its vast potential and the substantial security challenges it poses. Traditional access control methods, such as Role-Based Access Control (RBAC) and Discretionary Access Control (DAC), offer familiarity and flexibility but may struggle to address the complexities of big data. Innovative approaches like Attribute-Based Access Control (ABAC), data masking, and encryption provide more granular and adaptable solutions. Additionally, emerging technologies like homomorphic encryption promise advanced data privacy and security but come with computational overhead and adoption challenges.

Organizations must carefully evaluate their access control needs and the nature of their big data environment to determine the most suitable approach or combination of approaches. A holistic and well-balanced strategy that leverages both traditional and innovative methods, alongside emerging technologies, is essential to safeguard valuable data assets while enabling legitimate access and analysis. By addressing the unique access control challenges of big data, organizations can better protect sensitive information and mitigate the risks associated with cyber adversaries, ensuring the continued use of big data for informed decision-making and innovation.

Reference list

Aftab, M.U., Habib, M.A., Mehmood, N., Aslam, M. and Irfan, M. (2015). *Attributed role based access control model*. [online] IEEE Xplore. doi:https://doi.org/10.1109/CIACS.2015.7395571.

Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., Gehrke, J., Haas, L., Halevy, A., Han, J., Jagadish, H.V., Labrinidis, A., Madden, S., Papakonstantinou, Y., Patel, J., Ramakrishnan, R., Ross, K., Shahabi, C., Suciu, D. and Vaithyanathan, S. (2011). Challenges and Opportunities with Big Data 2011-1. *Cyber Center Technical Reports*. [online] Available at: https://docs.lib.purdue.edu/cctech/1/.

Bakshi, K. (2012). Considerations for big data: Architecture and approach. *2012 IEEE Aerospace Conference*. doi:https://doi.org/10.1109/aero.2012.6187357.

Balachandran, B.M. and Prasad, S. (2017). Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence. *Procedia Computer Science*, [online] 112, pp.1112–1122. doi:https://doi.org/10.1016/j.procs.2017.08.138.

Batko, K. and Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. *Journal of Big Data*, 9(1). doi:https://doi.org/10.1186/s40537-021-00553-4.

Cao, J. (2023). Security Management on Big Data of Business. pp.75–98. doi:https://doi.org/10.1007/978-981-99-3588-8\_5.

Casalicchio, E. and Iannucci, S. (2020). The state-of-the-art in container technologies: Application, orchestration and security. *Concurrency and Computation: Practice and Experience*, p.e5668. doi:https://doi.org/10.1002/cpe.5668.

Dattatray, M., Assistant, G. and Sulabha, A. (2014). Analysis of DAC MAC RBAC Access Control based Models for Security. *International Journal of Computer Applications*, [online] 104(5), pp.975–8887. Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=45a2775770d870b8675fb1301919224c9bcb7361.

Georgiadis, G. and 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, p.105640. doi:https://doi.org/10.1016/j.clsr.2021.105640.

Gupta, M., Patwa, F. and Sandhu, R. (2018). An Attribute-Based Access Control Model for Secure Big Data Processing in Hadoop Ecosystem. *Proceedings of the Third ACM Workshop on Attribute-Based Access Control*. doi:https://doi.org/10.1145/3180457.3180463.

Hossain, E., Khan, I., Un-Noor, F., Sikander, S.S. and Sunny, Md.S.H. (2019). Application of Big Data and Machine Learning in Smart Grid, and Associated Security Concerns: A Review. *IEEE Access*, [online] 7, pp.13960–13988. doi:https://doi.org/10.1109/access.2019.2894819.

Hu, V., Ferraiolo, D., Kuhn, R., Friedman, A., Lang, A., Cogdell, M., Schnitzer, A., Sandlin, K., Miller, R. and Scarfone, K. (n.d.). *NIST Special Publication 800-162 Guide to Attribute Based Access Control (ABAC) Definition and Considerations (Draft)*. [online] Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=cab698a5b0949aa7acd0858b55352c5df0a2c2fb.

Hu, V.C., Grance, T., Ferraiolo, D.F. and Kuhn, D.R. (2014). *An Access Control scheme for Big Data processing*. [online] IEEE Xplore. doi:https://doi.org/10.4108/icst.collaboratecom.2014.257649.

Koo, J., Kang, G. and Kim, Y.-G. (2020). Security and Privacy in Big Data Life Cycle: A Survey and Open Challenges. *Sustainability*, [online] 12(24), p.10571. doi:https://doi.org/10.3390/su122410571.

Lo Giudice, P., Musarella, L., Sofo, G. and Ursino, D. (2019). An approach to extracting complex knowledge patterns among concepts belonging to structured, semi-structured and unstructured sources in a data lake. *Information Sciences*, 478, pp.606–626. doi:https://doi.org/10.1016/j.ins.2018.11.052.

Marques, J. and Bernardino, J. (2020). Analysis of Data Anonymization Techniques. *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. [online] doi:https://doi.org/10.5220/0010142302350241.

Obaidat, M., Brown, J., Obeidat, S. and Rawashdeh, M. (2020a). A Hybrid Dynamic Encryption Scheme for Multi-Factor Verification: A Novel Paradigm for Remote Authentication. *Sensors*, 20(15), p.4212. doi:https://doi.org/10.3390/s20154212.

Obaidat, M., Brown, J., Obeidat, S. and Rawashdeh, M. (2020b). A Hybrid Dynamic Encryption Scheme for Multi-Factor Verification: A Novel Paradigm for Remote Authentication. *Sensors*, 20(15), p.4212. doi:https://doi.org/10.3390/s20154212.

OUTCHAKOUCHT, A., ES-SAMAALI, H. and Philippe, J. (2017). Dynamic Access Control Policy based on Blockchain and Machine Learning for the Internet of Things. *International Journal of Advanced Computer Science and Applications*, [online] 8(7). doi:https://doi.org/10.14569/ijacsa.2017.080757.

Rana, S.K., Rana, S.K., Nisar, K., Ag Ibrahim, A.A., Rana, A.K., Goyal, N. and Chawla, P. (2022). Blockchain Technology and Artificial Intelligence Based Decentralized Access Control Model to Enable Secure Interoperability for Healthcare. *Sustainability*, 14(15), p.9471. doi:https://doi.org/10.3390/su14159471.

Rawat, D.B., Doku, R. and Garuba, M. (2019). Cybersecurity in Big Data Era: From Securing Big Data to Data-Driven Security. *IEEE Transactions on Services Computing*, 14(6), pp.1–1. doi:https://doi.org/10.1109/tsc.2019.2907247.

Suraj, M.V., Kumar Singh, N. and Tomar, D.S. (2018). Big data Analytics of cyber attacks: a review. *2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA)*. doi:https://doi.org/10.1109/icscan.2018.8541263.