



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 9, September 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118



Blockchain Technology and it's Applications in Cybersecurity and Data Science: A Survey

Abhijeet Pachpute¹, Arbaz Attar², Shubham Narale³, Gaurav Durgude⁴

U.G. Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune

University, Pune, Maharashtra, India¹

U.G. Student, Department of IT Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune

University, Pune, Maharashtra, India²

U.G. Student, Department of IT Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune

University, Pune, Maharashtra, India³

U.G. Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune

University, Pune, Maharashtra, India⁴

ABSTRACT: Blockchain technology has been very effective in processing distributed transactions securely. there are many applications inclusive of dealing with cryptocurrencies such as bitcoin and smart contracts. Lately, the use of blockchain has been explored for data science applications. This paper surveys blockchain technology and discusses its applications in data science and cybersecurity. The high level of dependency on the internet and technology nowadays has led to new revenue streams and commercial enterprise models for businesses, but with this arises new gaps and opportunities for hackers to exploit. Due to blockchain distributed nature, blockchains offer no 'hackable' entrance or a critical factor of failure and thereby, provide more security as compared with numerous present database-backed transactional systems. The relationship between blockchain and data science, if there's any, has not been researched a lot. blockchain validates and data information, data science focuses on deriving meaningful insights from data for problem-solving. both of these technologies employ algorithms to control interactions with unique data segments.

KEYWORDS: Blockchain, Cybersecurity, Privacy, Data Science.

I. INTRODUCTION

Blockchain is a shared, well-established ledger that facilitates the process of noting transactions and tracking assets in a network. An asset can be physical (a house, car, cash, land) or nonphysical (intellectual property, patents, copyrights, branding). Effectively anything of value can be monitored and exchanged on a blockchain network, minimizing risk and reducing costs for everything involved. Information is the backbone of a business. The greater speed with which it is updated and received with high accuracy is always better. Blockchain is ideal for sharing that information as it provides real-time, sharing, and completely explicit information stored on an immutable ledger that can be retrieved only by authorized network members. A blockchain network can monitor payments, accounts, orders, production, and much more.

Blockchain Technology:

Blockchain consists of connection to different blocks forming chain of blocks. The data present in the block refers to the information regarding the transactions. The block also consists of three main elements data, nonce, and hash. A cryptographic hash is generated when the first block of the chain is created. Blockchains are known for their significant role in cryptocurrency systems, for maintaining a decentralized and secure record of transactions. The revolution with a blockchain is that it assures the constancy and security of a set of data and initiates trust without the need for a third party. One crucial difference between a normal database and a blockchain is how the data is structured. A set of information is held by the blockchain chain which collects the data in form of blocks. Blocks are closed and linked to previously closed blocks after a certain storage capacity of the block is filled, forming a chain of data known as the blockchain. A new block will be added for all the new information that is added to a certain block.



Blockchain for Cybersecurity:

Blockchain technology enables delegation through the involvement of members across a dispersed network. The system does not fail at any point and the user cannot block details. Even though, blockchain technologies are different in some crucial security aspects. Blockchain networks can be different in who can be involved and who has access to the data. The permissible and non-permissible are named as public and private networks respectively. They describe the authorization and accessibility of a network.

Blockchain for Data Science:

Everything is getting digitalized parallelly increasing the data that is generated on a daily basis. The data can be linked to any asset that is being used, right from the bank details to the IoT devices that operate the mechanisms of a house. Here, lots of sensitive data is generated. Such vital data is used for different analyses in the data science field. Data science produces very sensitive data that has a large impact on the institutions, such data needs to be kept securely. Data needs to be protected maintaining the data integrity while it is being transferred. Here, blockchain plays the crucial part of maintaining and holding the data securely. Blockchain maintains data integrity. While working in blockchain and data science go hand-in-hand, it makes it more sustainable and secure for data analysis.

II. RELATED WORK

Blockchain for Cybersecurity:

Blockchain technology adoption by companies and organizations will aid in the security of Internet data, information, devices, and services from cyberattacks. Blockchain carries responsibility for robust security against threats. The primary benefit of blockchain technology is that it integrates several computers, anonymous users, keys, and codes that are challenging to break. This section briefly discusses blockchain applications for security.

It is stated in [1] that blockchain is a technology that enables peer-to-peer (P2P) data transmission and storage. By using consensus-based algorithms, blockchain data could be consulted, shared, and protected structurally. It eliminates the need for intermediaries, or "trusted third parties," when utilized in a decentralized manner. Today, over 50 of the top publicly traded firms are exploring blockchain, and majority of them are willing to adopt it for the reliable basics. For example, in [1] it is stated that approximately 10 businesses, including FedEx, Burger King Russia, KIK, IBM, MasterCard, Microsoft, Overstock, Bank of America, Huawei Technologies, and Walmart have already adopted blockchain. The main objective of an organization's business continuity plan is to protect its assets from malicious attacks. In [2] we have discussed, that by using a blockchain-based automated resilience process, the different resilience process tasks may be coordinated effectively. Organizations may adopt a stronger resilience mechanism with the usage of blockchain, giving them the capacity to respond and recover fast from an attack. The risks related to the present methods are decreased by the usage of blockchain in the cyber resilience process.

Blockchain guarantees that the integrity of the backup is not compromised and stops backups from being saved without authorization. Additionally, because the data is kept across several nodes and the blockchain applies cryptography to secure its blocks, hacking a blockchain is difficult. The application offers a full history of each backup along with a timestamp. The hash value linked to each valid backup can be compared by the application to identify security vulnerabilities in the storage.

Since blockchain is decentralized, it guarantees the accuracy of data processing. It considerably aids in protecting system reliability from cyber threats. Because of the distributed and shared ledger nature of the blockchain, data visibility is enhanced, which can lead to higher security since it makes it easier to spot threats in their early stages and establish methods to rapidly eliminate them. The benefits of blockchain include enhancing verifiability to improve cyber security. We can ensure data is valid by having entities validate it both before and after a procedure. It is feasible to address the single point of failure and respond to circumstances in which attackers occupy several nodes when applying the blockchain to the national system for cyberdefence purposes. The blockchain data structure consists of a continuous hash chain. In the middle of the blocks, it is nearly impossible to remove or edit data. As a result, the participants can continue to audit the system data for cyber defence [3]. Blockchain provides decentralized smart contracts to enforce auditable multi-party payments based upon established standards among decentralized power suppliers and customers, commoditizing confidence in a step toward grid flexibility. Blockchain is essential to ensure secure data storage for the exchange of energy and the protection of privacy reported in [4]. Blockchain makes usage of the consensus protocol idea. These protocols are based on the idea of democratic systems and guarantee appropriate



network operation even in the case of malicious intrusions. Blockchain has been recognized as one of the most attractive technologies for a variety of activities, including cryptocurrency transactions, land records, healthcare, aviation, e-governance, data and network security, and many others, owing to the above-mentioned security attributes. Energy security and privacy preservation using blockchain technology descriptions can be found in [5].

Blockchain in Data Science:

Blockchain is a data structure that stores transactional records while also providing security, transparency, and decentralization. It is a structure that keeps public transactional records, also known as blocks, in many databases, known as the "chain," in a network linked by peer-to-peer nodes. This type of storage is sometimes referred to as a 'digital ledger.' It is a distributed ledger that is totally accessible to everybody on the network. Data Science, on the other hand, is a multidisciplinary field that extracts information and insights from structured and unstructured data using scientific methods, procedures, algorithms, and systems. It is a "concept to combine statistics, data analysis, machine learning, and their associated approaches" in order to use data to "understand and evaluate actual occurrences."

It is reported in [6] study of Blockchain technology, evaluating the issue in depth, and confirming the entirety of the data. Users depend on blockchain technology understanding to build a basic block network, evaluating input data provided through any file. The use of blockchain is merely to assure the data entirety, and the blockchain network is still simple with a few machines, allowing text removal to continue undetected. The development of blockchain technology is broad, such as the creation of a system applicable to agencies and organizations in issues assuring that the agency's materials are intact and utilized on the complete computer network to jointly participate on the network to improve the entirety and transparency for documentation. Each blockchain solution, however, employs a chain structure as well as innovative data structures. Clearly, this data may be studied to give new insights on upcoming patterns. Blockchain technology has lately sparked consumer and industry interest in a wide range of applications, including digital money, food safety, health care, and weapon monitoring. As additional Blockchain applications emerge on a daily basis, so does the complexity and volume of data held by Blockchain, resulting in a new separate study path of Blockchain Data Analytics. A concise introduction of the present status of Blockchain Data Analytics, with emphasis on methodological development and new research difficulties, as well as insight into some of the most relevant financial applications are discussed in [7].

According to [8], work addresses the data science skills gap with a Blockchain-driven method to connecting data science education to industry. Smart Blockchain Badges are used to assist learners in achieving their professional objectives. In the coming phases, implementation of Smart Blockchain Badges for usage by actual learners and assess their use for lifetime learning. It will also expand the work outside the realm of data science by implementing it in popular general-purpose online learning platforms like the FutureLearn12 MOOC platform and the OpenLearn13 OER repository. It is also stated in [8], how Smart Blockchain Badges may help learners progress their careers in data science by providing personalized recommendations based on their learning accomplishments. This effort intends to improve data science accreditation by implementing a strong system based on Blockchain technology. Learners will benefit from an advanced, open, and transparent accreditation system, as well as employment suggestions that match their talents and have the ability to advance their careers. Smart Blockchain Badges provide considerable benefits to learners by assisting them in structuring their earning path to meet their professional objectives. Through the matching of badges to job profiles, Smart Blockchain Badges may be used to connect accreditations to employment. Smart Blockchain Badges may be used to associate acquired badges with organisational positions in order to promote internal employees. As a result, by connecting data science education to business, this effort contributes to addressing the data science skills gap.

While Data Science focuses on leveraging data to improve administration, blockchain's distributed ledger ensures data security. These technologies hold a lot of untapped potential for increasing efficiency and production. The threat of privacy and data security pervades every step in the data science industry chain, including data production, gathering, processing, and sharing, and the reasons of risk are complex and intertwined. Blockchain technology has received great acclaim and recognition for its decentralized architecture, anonymity, security, and other features. It will alter how user get and exchange information. Users think that blockchain technology has the potential to overcome several restrictions in data science and boost its advancement. Data science and blockchain technology may be used together to significantly transform how to process and interpret data. It begins by providing an overview of data science, covering



big data technologies and the risks it confronts. Then it introduces blockchain technology, including blockchain architecture and certain major blockchain characteristics. It also illustrates the possible benefits, limitations, and unresolved concerns of integrating blockchain with data science. While embracing blockchain technology, it should increase blockchain technology guidance and standards, focus on data security risk research and analysis, continuously monitor development patterns, and proactively explore regulatory methods description can be found in [9].

Big data and blockchain are being praised as the next big things that will transform the way businesses operate. The majority of people feel that these technologies are mutually incompatible, with each having its own path to independent use. Data Scientists are now leveraging Blockchain technology to verify data validity and tracking at every stage along the chain. One of the reasons for its widespread adoption is its unchangeable security. Data is safeguarded at every stage by Blockchain's decentralized ledger, which uses multiple signatures. Blockchain enables real-time analysis, data traceability, maintains data accuracy, facilitates data sharing, fosters trust, and improves data integrity in data science. Lastly in [10] Blockchain technology research and applications in data science are investigated. The expansion of social information and networks has resulted in an explosion of data in today's society. As the number and diversity of data has grown, researchers have been urged to make business judgments by analysing big data. This has also accelerated the development of the data science sector. Every component in the data science industry chain is subject to data security and privacy concerns, including data collection, storage, processing, and sharing, and the reasons of risk are complex and interrelated. Blockchain technology is highly lauded and recognized for its decentralized design, anonymity, security, and other properties. It will revolutionize the way we collect and transmit information. We believe that blockchain technology has the potential to address several shortcomings in data science and promote its growth.

III. CONCLUSION

In this paper, we surveyed and discussed blockchain technology and its applications in data science and cyber security. Especially, we discussed various concepts such as blocks, transaction execution, bitcoin, smart contracts, cryptographic checksum, and associated blockchain components and then discussed how blockchain technology may be used for data science such as data analytics and data sharing. Blockchain offers protection for the entire data lifecycle. Subsequently, the applications of blockchains for security along with IOT security, storage, DDOS attacks, confidentiality, and integrity. In addition to authentication, the paper examined various aspects of data science activities including the privacy-conscious policy-based data life cycle system and discover how blockchain technology can be securely implemented for various distributed transactions involved in those activities. However, there are some areas that can be explored thoroughly such as data-backed financial transactions in data science and ransomware attacks in terms of cybersecurity.

REFERENCES

- [1] A. Farion, O. Dluhopolskyi, S. Banakh, N. Moskaliuk, M. Farion and Y. Ivashuk, "Using blockchain Technology for Boost Cyber Security," 2019 9th International Conference on Advanced Computer Information Technologies (ACIT), 2019, pp. 452-455, doi: 10.1109/ACITT.2019.8780019.
- [2] T. Harman, P. Mahadevan, K. Mukherjee, P. Chandrashekar, S. Venkiteswaran and S. Mukherjee, "Cyber Resiliency Automation Using Blockchain," 2019 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), 2019, pp. 51-54, doi: 10.1109/CCEM48484.2019.00011.
- [3] S. Lee and S. Kim, "Blockchain as a Cyber Defense: Opportunities, Applications, and Challenges," in IEEE Access, vol. 10, pp. 2602-2618, 2022, doi: 10.1109/ACCESS.2021.3136328.
- [4] G. Chen, M. He, J. Gao, C. Liu, Y. Yin and Q. Li, "Blockchain-Based Cyber Security and Advanced Distribution in Smart Grid," 2021 IEEE 4th International Conference on Electronics Technology (ICET), 2021, pp. 1077-1080, doi: 10.1109/ICET51757.2021.9451130.
- [5] K. Kaur, G. Kaddoum and S. Zeadally, "Blockchain-Based Cyber-Physical Security for Electrical Vehicle Aided Smart Grid Ecosystem," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 8, pp. 5178-5189, Aug. 2021, doi: 10.1109/TITS.2021.3068092.
- [6] Khuat Thanh Son, Nguyen Truong Thang, Tran Manh Dong, Nguyen Ha Thanh, Blockchain Technology for Data Entirety, Science Research. Volume 6, Issue 6, December 2018, pp. 68-75. doi: 10.11648/j.sr.20180606.12
- [7] Dixon, Matthew & Gel, Yulia & Kantarcioglu, Murat & Akcora, Cuneyt. (2018). Blockchain Data Analytics. IEEE Intelligent Systems. Blockchain Data Analytics, Article in IEEE Intelligent Systems, December 2018



- [8] A. Mikroyannidis, J. Domingue, M. Bachler and K. Quick, "Smart Blockchain Badges for Data Science Education," 2018 IEEE Frontiers in Education Conference (FIE), 2018, pp. 1-5, doi: 10.1109/FIE.2018.8659012.
- [9] Liu, Jiameng & Peng, Shaoliang & Long, Chengnian & Wei, Lijun & Liu, Ophelia & Tian, Zhihui. (2020). Blockchain for Data Science. 24-28. 10.1145/3390566.3391681.
- [10] Ayyappan, Vasanthi. (2021). Blockchain for Data Science and its Business Applications.



INNO  SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN  INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details