DATA MANAGEMENT AND SECURITY IN BLOCKCHAIN SYSTEMS

Editors:
Sonali Vyas
Shaurya Gupta
Vinod Kumar Shukla
Chinwe Peace Igri

Bentham Books

Data Management and Security in Blockchain Systems

Edited by

Sonali Vyas
School of Computer Science
UPES Dehradun, Uttarakhand, India

Shaurya Gupta
School of Computer Science
UPES Dehradun, Uttarakhand, India

Vinod Kumar Shukla
Department of Engineering and Architecture
Amity University Dubai, UAE

&

Chinwe Peace Igri
Computer Science and Mathematics
Mountain Top University
Port Harcourt, Rivers State, Nigeria

Data Management and Security in Blockchain Systems

Editors: Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla and Chinwe Peace Igri

ISBN (Online): 978-981-5305-81-4

ISBN (Print): 978-981-5305-82-1

ISBN (Paperback): 978-981-5305-83-8

First published in 2024.

need for a court order if at any point you breach any terms of this License Agreement. In no event will any delay or failure by Bentham Science Publishers in enforcing your compliance with this License Agreement constitute a waiver of any of its rights.

3. You acknowledge that you have read this License Agreement, and agree to be bound by its terms and conditions. To the extent that any other terms and conditions presented on any website of Bentham Science Publishers conflict with, or are inconsistent with, the terms and conditions set out in this License Agreement, you acknowledge that the terms and conditions set out in this License Agreement shall prevail.

Bentham Science Publishers Pte. Ltd.
80 Robinson Road #02-00
Singapore 068898
Singapore
Email: subscriptions@benthamscience.net

CONTENTS

PREFACE

In an era of digital transformation, Blockchain technology, the Internet of Things (IoT), Artificial Intelligence (AI), and big data systems have ushered in a new paradigm of data management and security challenges. As organizations embrace the potential of these disruptive technologies to revolutionize various industries, the need for robust data management practices and stringent security measures becomes increasingly paramount.

This book delves into the intricate landscape of data management and security within blockchain-based systems, exploring the multifaceted dimensions of this evolving field. A comprehensive evaluation examines the fundamental principles and practices governing data management in Blockchain networks, addressing the complexities inherent in ensuring data integrity, confidentiality, and availability.

Central to this discourse is the symbiotic relationship between Blockchain technology and IoT, as they collaborate to fortify data security and streamline traffic management. By leveraging blockchain's immutable and decentralized nature, coupled with the connectivity and sensor capabilities of IoT devices, novel solutions emerge to mitigate security risks and optimize data handling processes.

Furthermore, this book elucidates the security challenges confronting Blockchain networks, elucidating the evolving threat landscape and vulnerabilities inherent in these decentralized systems. From cyber-attacks targeting big data repositories to the vulnerabilities plaguing IoT devices, each chapter dissects the intricacies of modern-day data security threats and proposes innovative solutions to fortify system resilience.

Moreover, it explores integrating AI-enabled security mechanisms within Blockchain systems, harnessing the power of machine learning and predictive analytics to identify and thwart potential threats proactively. Organizations can use this synergy to elevate their defense mechanisms, preemptively addressing security breaches and safeguarding critical data assets.

Finally, this book endeavors to serve as a comprehensive guide for practitioners, researchers, and policymakers grappling with data management and security complexities in the digital age. By offering insights into emerging trends, best practices, and technological advancements, it aims to empower stakeholders to navigate the intricate landscape of blockchain-based systems with confidence and resilience.

As we embark on this journey through data management and security in Blockchain networks, let us unravel the intricacies, confront the challenges, and embrace the transformative potential of these groundbreaking technologies.

Sonali Vyas
School of Computer Science
UPES Dehradun, Uttarakhand, India

Shaurya Gupta
School of Computer Science
UPES Dehradun, Uttarakhand, India

Vinod Kumar Shukla
Department of Engineering and Architecture
Amity University Dubai, UAE

&

Chinwe Peace Igri
Computer Science and Mathematics
Mountain Top University
Port Harcourt, Rivers State, Nigeria

List of Contributors

Ashish Tiwari	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India
Anusha Viswanadapalli	Department of Computer Science & Engg., VFSTR University, Andhra Pradesh, India
Ashima Bhatnagar Bhatia	Vivekananda Institute of Professional Studies-TC, New Delhi, India
Bhakti Thakre	Computer Science Engineering (Cyber Security), St. Vincent Pallotti College of Engineering and Technology, Nagpur, India
Bhanu Prakash Lohani	Amity University, Greater Noida, UP, India
Bhuvi Sharma	Amity University, Greater Noida, UP, India
Bhairab Sarma	Department of Computer Science, University of Science and Technology, Ri-Bhoi, Meghalaya, India
Deepshikha Bhargava	Amity University, Greater Noida, UP, India
Jhansi Bharathi Madavarapu	University of the Cumberlands, Williamsburg, KY 40769, USA
Kavita Sharma	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India
Khushboo Jain	Department of Computer Science & Engg., Indian Institute of Information Technology, Nagpur, India
Khushi Dadhich	Amity University, Greater Noida, UP, India
Lipsa Das	Amity University, Greater Noida, UP, India
M.M. Dhabu	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India
Malla Abhilash	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India
Meera Dhabu	Department of Computer Science & Engg., Visvesvaraya National Institute of Technology, Nagpur, India
Nitesh A. Funde	Department of AI, SardarVallabbhai National Institute of Technology (SVNIT), Surat, Gujarat, India
Pallabi Baruah	Department of Computer Science, University of Science and Technology, Ri-Bhoi, Meghalaya, India
Praveen Nelapati	School of Computer Science & Engg., VIT-AP University, Amaravati, Andhra Pradesh, India
Pawan Whig	Vivekananda Institute of Professional Studies-TC, New Delhi, India
Ramya Thatikonda	University of the Cumberlands, Williamsburg, KY 40769, USA
Suman Avdhesh Yadav	Amity University, Greater Noida, UP, India

Uma Yadav	Computer Science and Engineering (Data Science), Shri. Ramdeobaba College of Engineering and Management, Nagpur, India
V.S.S. Koushik	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India
Vrudhula Sreedhar	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India
V.V. Jithin	Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India

CHAPTER 1

Evaluation of Data Management in Blockchain-based Systems

Bhakti Thakre[1,*] and **Uma Yadav**[2]

[1] *Computer Science Engineering (Cyber Security), St. Vincent Pallotti College of Engineering and Technology, Nagpur, India*

[2] *Computer Science and Engineering (Data Science), Shri. Ramdeobaba College of Engineering and Management, Nagpur, India*

Abstract: Blockchain records every data transaction on its network through a distributed digital ledger that is accessible to the public. The agreement-based process of recording and updating data across dispersed nodes is crucial for enabling trustless multi-party transactions in blockchain-based systems. The degree of utility and performance of a blockchain-based application is ultimately determined by understanding what and how the data is stored and changed. By offering an immutable and consistent data storage technology, it improves the quality of the data while posing new data management issues.

It analyzes blockchains from the viewpoint of a developer to highlight important concepts and considerations when incorporating a blockchain into a larger software system as a data store. Data Management involves architectural layers for storing data and conceptualizing each layer in blockchain, examining the flow of data in blockchain-based applications, andexploring data administration aspects for blockchains. Data domination issues in blockchains are related to privacy and Quality Assurance (QA). The privacy of data can be preserved by keeping it in an encrypted form, but it affects usability and flexibility in terms of effective search. Attribute-based Searchable Encryption (ABSE) has proven its worth by providing fine-grained searching capabilities in the shared cloud storage.

In order to emphasize key ideas and things to keep in mind when integrating a blockchain as a data storage system into a larger software system, it analyzes blockchains from the perspective of a developer. Data management includes creating architectural layers for data storage, conceptualizing each layer in a blockchain, analyzing data flow in blockchain-based applications, and finally investigating data administration features for blockchains. The problems with data dominance in blockchains concern Quality Assurance (QA) and privacy. Data privacy can be maintained by encrypting it, but this compromises flexibility and usability in terms of efficient search. Since it allows for more precise searching in shared cloud storage, attribute-based searchable encryption, or ABSE, has shown its value.

[*] **Corresponding author Bhakti Thakre:** Computer Science Engineering (Cyber Security), St. Vincent Pallotti College of Engineering and Technology, Nagpur, India; E-mail: bthakre11@gmail.com

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

The vulnerability of cloud services to assaults stems from their widespread accessibility. In cloud computing, data tampering is a risk to data integrity that can happen. Clients using cloud computing across a range of application areas demand assurances regarding the veracity and accuracy of their data.

Keywords: ABSE, Block chain, Cloud computing, Data integrity, Encryption, Fine-grained, QA.

INTRODUCTION

The potential for blockchain technology to revolutionize society has been likened to that of the WWW. The foundations of blockchain have been supported by a wide number of other applications in a short period of time, beyond the original purpose of cryptocurrency. These applications include asset management, insurance, finance, and medical/health. From the standpoint of these applications, blockchain enhances data quality by providing transparency, immutability, and consistency [1].

The architecture of block-chains, which provides these benefits, however, introduces current issues with data-management. As an ex, the following problems with blockchain as a network for data processing and storage may be identified as unresolved: Blockchain data formats include document and key-value store, which are frequently integrated with “off-chain” data storage. Therefore, ad hoc and handmade programming are needed in BC-based structures to search for and repossess a variety of data, unlike abstract and declarative query strategies in conventional databases. Sympathetically, how to get, assimilate, and analyze data in this assorted situation is crucial, especially in light of the growing need for large-scale blockchain data analytics.

The amount of information stored and managed by blockchain networks will only increase over time. However, many modern systems have low throughput, scalability, and latency. Moreover, open block-chains charge subscriptions for storing and updating information to deter idle data. Some of these issues can be resolved by carefully analyzing the on-chain and off-chain information structural adoptions completed by a block chain use.

Blockchain technology provides permanent and network-wide access to data storage. Concerns about quality and privacy, among other data governance issues, are brought up by this. Encrypting data is recommended, but doing so could leave it vulnerable to future brute-force decryption attacks (quantum computing advancements, for instance, could make current encryption techniques ineffective) or cause unintentional privacy breaches. The development of proper frameworks

for blockchain data governance is therefore necessary in order to support effective management and responsible application of BC expertise.

It is imperative to investigate the usage of a digital ledge as an information storage medium in the framework of information organization, given these difficulties. Wherever there is a digital ledger, database managers and application developers can create and oversee a large software framework. In order to coexist with greater effectiveness, an auxiliary database must have a thorough understanding of blockchain technology, specifically with regard to data handling and storage. Errors, poor designs, and issues resulting from false presumptions about how block chains work should also be avoided. In other works, the functionality and distinctive characteristics of blockchain have been briefly compared to those of databases [2-5]. An addition to these efforts, we further construct the distinctions according to the way layers of software systems are commonly viewed by application developers.

The blockchain technology is methodically examined in this study from a database perspective. With the goal of improving the usefulness and appropriate usage of block chains in big software systems, we want to comprehend block chains as data storage better. To do this, we pinpoint and examine key data management challenges in the growth and administration of digital ledge-based systems. The subsequent contributions are:

- Suggest a novel understanding of blockchain as the data repository for an application.
- Identify and assess the most effective approaches to operational problems and blockchain data structures.
- Examine the data management elements of block chains.
- Provide relevant, actionable insights into the rapidly developing fields of digital ledge analytics of data and reliable blockchain-based data examine.
- Examine the administration of blockchain records confidentiality and superiority, including existing problems and potential future possibilities.

BLOCKCHAIN PERFORMANCE

Blockchains can offer a reliable and impartial data storage policy aimed at a sizable system that incorporates the technology. Trust and neutrality arise from the following features of the system, consensus mechanism, and cryptographic procedures it uses, along with the unique architecture of the ledger structure:

CHAPTER 2

Data Security and Traffic Management Using Iot and Blockchain Application

Lipsa Das[1,*], Bhanu Prakash Lohani[1], Deepshikha Bhargava[1] and **Bhuvi Sharma[1]**

[1] *Amity University, Greater Noida, UP, India*

Abstract: The integration of Internet of Things (IoT) and blockchain technologies offers promising solutions for ensuring robust data security and efficient traffic management in contemporary urban environments. This book chapter explores the synergy between IoT and blockchain applications, presenting novel approaches to address the evolving challenges of data integrity, privacy, and traffic congestion. Through the convergence of these technologies, innovative mechanisms for securely collecting, transmitting, and storing data from IoT devices are introduced, fostering trust and transparency in data transactions. Additionally, leveraging blockchain's decentralized ledger, smart contracts, and cryptographic principles, the chapter elaborates on how immutable records can streamline traffic management systems, enabling real-time monitoring, optimization, and enforcement of traffic regulations. This chapter elucidates the potential of IoT and blockchain integration to revolutionize data security and traffic management paradigms, paving the way for smarter, safer, and more sustainable urban ecosystems.

Keywords: Blockchain, Data security, IoT, Traffic management.

INTRODUCTION TO SMART CITIES

"Smart city" refers to modern metropolitan regions that use ICT (Information and Communication Technology) to increase production, communicate information with citizens, and ensure higher-quality public services. Using cutting-edge technologies, its main objective is to improve city services and accelerate economic growth while also ameliorating inhabitants' quality of life (IoT, AI, Blockchain, *etc*). Instead of counting the number of pieces of technology used, cities are rated according to how they use it [1].

[*] **Corresponding author Lipsa Das:** Amity University, Greater Noida, UP, India; E-mail: lipsaentc9@gmail.com

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

The following characteristics characterize a smart city:

- Technology based on infrastructure.
- E-governance: It entails taking part in government planning and decision-making processes, enhancing democratic laws and public & social services, and preserving transparency between the government and the people.
- Mobility in cities (Transportation).
- Smart living (education and cultural facilities, healthcare): It entails novel concepts and inventions that guarantee a higher standard of living for people—things like efficiency, economy, sustainability, and productivity, among others. Additionally, it includes providing each citizen access to quality, affordable healthcare and education [2].
- Intelligent environment (initiatives towards nature, pollution, Sustainable resource management).

In addition to the technologies used, data analytics is crucial to the growth of smart cities. Here, data analysts evaluate the information provided by smart city technologies and, if necessary, make enhancements.

New developments like automation, AI, ML, and the IoT are encouraging the development of smart cities. It also uses a mix of software, User Interfaces (UI), Internet of Things (IoT) gadgets, and communication-based networks.

However, there is significant reliance on the IoT, a network interconnecting various objects such as cars, sensors, and home appliances. This IoT network facilitates the exchange of data and communication among these interconnected entities. Sensors and devices within the IoT collect and transmit data, subsequently stored on servers or in the cloud. The integration of Data Analytics (DA) and the networking capabilities of these devices bring about the convergence of physical and digital components within cities. This convergence enhances the efficiency of both public and commercial sectors, leading to economic benefits and an overall enhanced quality of life for individuals.

From a logical standpoint, any aspect of municipal management can be integrated into a smart city initiative. A case in point is the implementation of smart parking meters, wherein an application aids vehicles in locating available parking spaces without navigating congested city blocks unnecessarily. The smart meter also facilitates digital payments, eliminating concerns about the meter running out of funds. Within the transportation sector, smart traffic management monitors and analyses traffic patterns to optimize street lighting and prevent excessive congestion during peak hours. Smart public transit is another integral component of smart cities, with transit systems employing real-time service synchronization to enhance efficiency and user satisfaction. Bike and vehicle-sharing services are

also prevalent in smart cities, contributing to increased productivity and overall satisfaction.

Smart cities utilize their network of IoT devices and other connected technology to boost liveability and encourage economic development [3]. Successful smart cities follow four steps:

- **Collection:** Real-time data collection is performed by smart sensors placed all around the city.
- **Analysis:** The information acquired by the smart sensors is assessed to produce insightful conclusions.
- **Communication:** Through efficient networks of communication, decision-makers are informed of the insights gleaned from the analytical phase.
- **Take Action:** Cities make use of the data's information to create solutions, enhance operations and asset management, and increase resident quality of life.

IOT ENABLED SMART CITIES

IoT is a network of heavily networked, individually identifiable equipment, objects, and other items. Data can be transferred between these devices over a network. IoT gadgets can sense, gather, and transmit information *via* the Internet in a straightforward manner. Any society's progress depends heavily on transportation and logistics. Over the past ten years, the academic and practice communities have shown a lot of interest in the IoT. Specific OS and interaction-based protocols are required to make communications between humans and gadgets possible in order to achieve crucial IoT features [4].

In the application field, Fig. (**1**) demonstrates a broad overview of the Internet of Things. Due to growth in information and communications technology, cities in the new era have become “smarter” and more productive in many ways (ICT). However, not all of the components in smart cities need to be considered “smart” because this is not required nor even practicable. The price and accessibility of the necessary technologies have a considerable impact on how quickly the smart components in Fig. (**1**) will be deployed [1]. IoT implementation in various application areas has been the focus of various projects in Europe funded by the 7th Framework Program for Research and Technological Development (FP7), such as District Information Modeling and Management for Energy Reduction (DIMMER) [5], Advancing Identification Matters (AIM) [6], Smart Energy Efficient Middleware for Public Spaces (SEEMPubS) [7] and Intelligent Use of Building Energy Information (IntUBE) [8]. To date, smart cities have used communication and network technology to convey contemporary town problems like overpopulation, crowdedness, and traffic jams [9].

CHAPTER 3

Data Management and Security in Blockchain Networks

M.M. Dhabu[1], Ashish Tiwari[1], Kavita Sharma[1,*], V.S.S. Koushik[1], Vrudhula Sreedhar[1], V.V. Jithin[1] and **Malla Abhilash[1]**

[1] *Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India*

Abstract: Blockchain is one of the major enhancements in this era of information technology. In this chapter, we will dig in deep for data management and security in blockchain networks. Blockchain is a decentralized, immutable, traceable, and trustworthy network for storing and retrieving information. As far as the need to store data is concerned, various factors come into consideration, such as network security, data security, cyber security risk, and data integrity in storage systems. Storing data in blockchain gives us rapid query processing and enhanced data auditing schemes with its public ledger, which leads to low computation costs and quick traceability. Here, data is distributed and decentralized. Storing data in chains leads to better data quality, owing to block chains' immutability, transparency, and traceability, which enables users to get better analytics and mining results. Generally, the main threats to data are users' privacy, anonymity, and data tampering. Storing data in the blockchain ensures these issues don't arise. Being the whole process is auditable, consistency and accuracy of blockchain are not compromised, although it gives high safety and security markups to the data stored. Current implementations of blockchain in real time show high computational cost, time, and latency, which leads to low scalability. High levies are charged to users for peer-to-peer transactions in public blockchains. However, the most discussed issues are related to blockchain security, data management challenges, and policies around the world.

Keywords: ADS, Byzantine fault tolerance, Cross-chain, Cosmos, ChainSQL, Distributed merkle tree, Hybrid blockchains, LSM trees, Merkle b-tree, Omniledger, Penetration testing, Polkadot, Ripple, Sharding, Side-chains, Tendermint.

INTRODUCTION

Numerous blockchain projects, including Ethereum, bit-coin, and hyper ledger, have attracted a lot of attention from both academic perspectives and business use

[*] **Corresponding author Kavita Sharma:** Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, 440010, India; E-mail: kavitamahy14feb@gmail.com

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

cases. Common characteristics of blockchains include decomposition, persistency, obscurity, and audibility. Blockchain systems are divided into three types on the basis of their chain structures: standard blockchains, DAG-based blockchains, and hybrid blockchains [1].

Blockchain is a decentralized database system that consists of several numbers of nodes distributed across networks. The records are immutable *i.e.*, they cannot be altered, ensuring the integrity of the data. The decentralized nature of blockchain enhances its security as compared to centralized systems. Transactions are processed through consensus mechanisms, reducing the chances of fraud. A blockchain allows for transparency, as all parties in the network have access to the same information. The decentralization also ensures that no single entity has control over the network. Blockchain technology is widely used in crypto-currencies but has potential applications in various industries.

Users of blockchain networks can identify if there is an inconsistency in data because all blocks are linked to each other. The hash value associated with the block will get modified if its content is modified, and as a result, all ensuing blocks' hash values will likewise modify. The network users rapidly notice the inconsistent state this causes in the blockchain. After that, they will reject the modified block and continue to work with the unaltered, original blockchain. The immutability and integrity of the data recorded on the blockchain are therefore guaranteed.

Blockchain technology, while regarded as a secure solution, still poses several security challenges that must be addressed. One of the issues is the potential occurrence of a 51% attack, whereby an adversary could attain control over more than half of the network's computing power, thus enabling manipulation of the blockchain. Another security concern is the compromise of private keys, which, if lost or stolen, results in permanent loss of funds. Inadequately written or vulnerable smart contracts may also result in security breaches. Hackers often employ phishing and social engineering techniques to deceive individuals into revealing their private keys. The decentralized and anonymous nature of blockchain may also raise regulatory issues, resulting in restrictions or limitations on its usage.

The scalability and data load of a blockchain can have a profound impact on its performance. As the size of a blockchain network increases, its complexity also increases, which affects the system's throughput.The use of consensus mechanisms, such as proof of work, can slow down transaction processing, leading to increased latency and decreased user experience. The decentralized nature of the network requires all participants to validate transactions,

contributing to longer processing times. The increased number of nodes in the network can also result in increased network overhead and decreased performance [2]. The growing amount of data stored on the blockchain also increases storage requirements for each node, raising costs. Off-chain solutions and sharding can improve scalability and performance but may also introduce new security risks that must be evaluated [3].

BACKGROUND

Traditionally, offline word processors like Microsoft Word were used for sharing documents from one person to another, where one would create the document and send it to others for modification to write their own ideas. However, two persons can not update the document simultaneously. Then simultaneously, editable online word processors like Google Documents came into existence wherein anyone can create the document and give access rights to the selected person. However, here, the problem is that the environment is still centralized. The problem with a centralized system is that it does not handle a single point of failure. If you don't have sufficient bandwidth to load the document, you will not be able to edit. What if the server crashes? All the data are stored at a single point on the server. If it crashes, users would be the victims as they have no idea how to retrieve the documents, and if the server is subject to any cyber-attack that results in data and privacy risk, the user again bears the brunt and is at a loss without any of their mistakes. Single points of failure, lack of transparency, vulnerability to cyber-attacks, limited scalability, unequal distribution of power, inefficiency, and users being prone to censorship are some of the disadvantages of centralized systems. In a decentralized system, we need to coordinate with multiple points, which is time-consuming, and in a distributed system, everyone collectively executes the job at the same time. The plausible solution is that everyone can edit on their local copy of the document, and the internet will take care of ensuring consistency.

Three popular blockchain designs manage distributed ledgers using various topologies. The hybrid blockchain handles multiple chains, while the regular blockchain uses the list chain structure. A DAG-based tangle is used in blockchains. The way a blockchain network arranges transactions is referred to as a blockchain data structure.

DATA MANAGEMENT TECHNIQUES IN STANDARD BLOCKCHAIN

Blocks are connected to each other in a linear manner in a standard blockchain. There are a number of transactions present inside a block. Blocks are made up of block headers and block bodies (which contain transaction data). These blocks also store the details about the transactions (known as block metadata) in the

CHAPTER 4

Data Management, Security Challenges, and Solutions in Blockchain Network

Pallabi Baruah[1,*] and **Bhairab Sarma**[1]

[1] *Department of Computer Science, University of Science and Technology, Ri-Bhoi, Meghalaya, India*

Abstract: Data management challenges are one of the primary concerns in business-related organizations. It may result in poor risk management decisions, data loss, data breaches, illegal access, data silos, noncompliance with legislation, an unregulated environment, a limited number of resources, and so on. Again, data security, or information security, includes the practices, policies, and principles to protect digital data and other kinds of information. Data security is based on three foundational principles — confidentiality, integrity, and availability. All operations around the globe are becoming increasingly reliant on data to operate their day-to-day operations and make educated business decisions. With so much data being created, it has become challenging to manage data throughout the enterprise, which may be dispersed over different geolocations and using tens of business line applications. It's accepted that data is the new oil for any organization. There may be very crucial data that may result in tremendous opportunities for companies, which is why it is vital to have well-defined data management plans in place in order to face the most difficult challenges that data management implies. In this chapter, emphasis is given to the most common data management challenges and pain points and to solving them in blockchain networks. It also includes the implementation of blockchain network technology in a wide range of applications, opportunities, and challenges.

Keywords: Blockchain, Challenges, Decisions, Management, Opportunities, Risk.

INTRODUCTION

Blockchain technology is a mechanism that is advanced and allows sharing of information to be transparent within a set of business networks. The concept of blockchain plays an important role in cryptocurrency for maintaining a secure and decentralized record of transactions. The data are stored in blocks in a chain. So,

[*] **Corresponding author Pallabi Baruah:** Department of Computer Science, University of Science and Technology, Ri-Bhoi, Meghalaya, India; E-mail: pallavibaruahguha@gmail.com

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

data is consistent as the chain cannot be deleted or modified without the proper consent. Therefore, an unalterable or immutable ledger is used to track orders, payments, accounts, and transactions. A blockchain is distributed in the sense that multiple copies are saved on many machines, and they must all match for it to be genuine [1].

The blockchain collects transaction information and enters it into a block, similar to a cell in a spreadsheet containing information. This information is processed through an encryption algorithm, which generates a hexadecimal number called the hash when the block is full.

HOW DOES BLOCKCHAIN TECHNOLOGY WORK?

Transactions are recorded as blocks of data. Who, what, when, where, how much, and even the conditions are recorded by the block of data. Those transactions show the movement of an asset that can be tangible or intangible.

Each block is connected to the previous and next ones. These blocks form a chain of data as an asset moves from place to place or ownership changes hands. Again, the blocks are connected securely together to avoid any block from being changed or a block being added in between two existing blocks. The blocks contain the exact time and sequence of transactions [2].

Transactions are blocked together in an irreversible chain: a blockchain. Each additional block maintains the verification of the previous block and, hence, the entire blockchain, which is depicted in Fig. (**1**). This renders the blockchain tamper-evident, delivering the key strength of immutability. This helps to avoid the possibility of tampering by a malicious user and builds a ledger of transactions and other network members that is trustworthy [3].

Distributed Database: There is no Central Server or System that keeps the data of the Blockchain. The data is distributed over millions of computers around the world, and they are connected to the blockchain. This system allows the Notarization of Data as it is present on every Node and is publicly verifiable [4].

BENEFITS OF BLOCKCHAIN TECHNOLOGY

Blockchain has achieved enormous popularity worldwide; therefore, companies and individuals see clear benefits from technology over traditional systems. It's a technology that is incredibly diversified. It is so diverse that it is applicable in almost any business industry today. Some of these include banking, healthcare, e-commerce, mining, logistics, and transport. The list is never-ending [5].

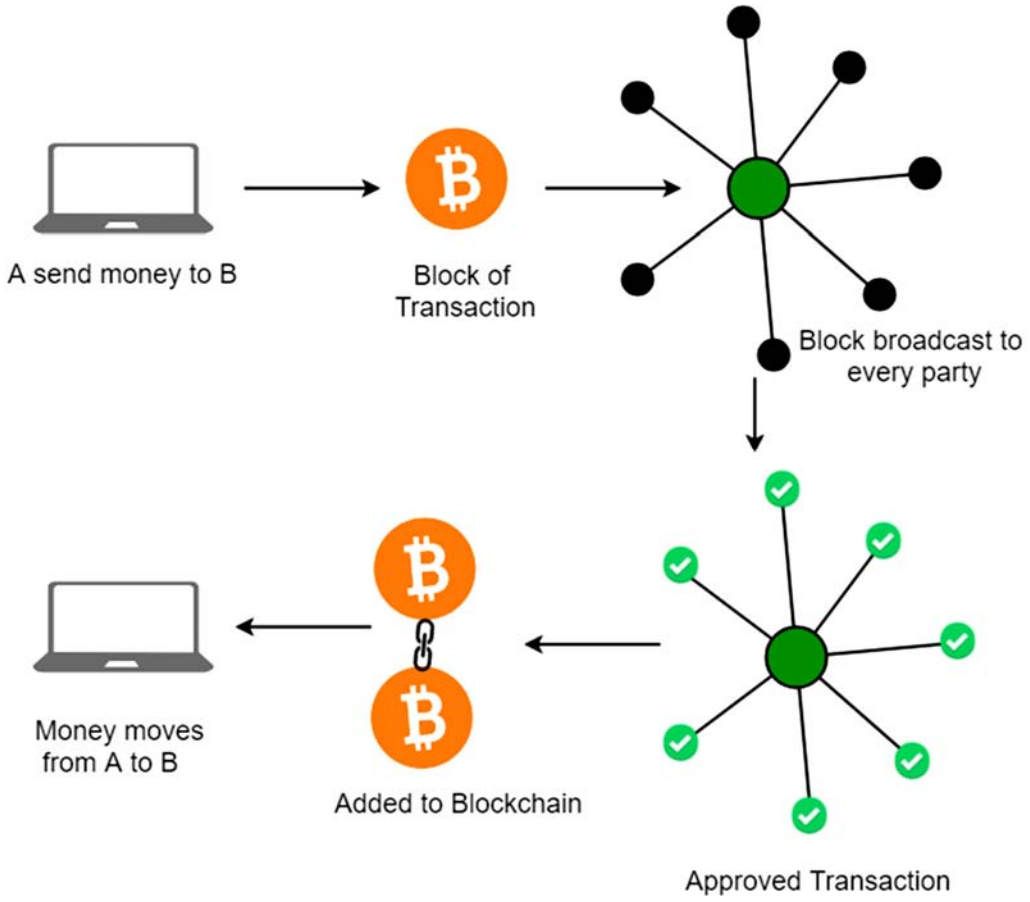

Fig. (1). Blockchain Technology.

Decentralised Structure

Blockchain makes it possible to share data within a vast ecosystem where no single organization has sole control, in addition to enabling transparent transactions. A related instance is the supply chain: Many firms- from suppliers and transportation providers to producers, distributors, and retailers- demand or require information from other businesses in that chain, but no one is responsible for facilitating all that information exchange. This issue is maintained by Blockchain because of its decentralized structure.

Transparency

Anyone who is in the network can see the change as well as the updated record when a transaction history is updated. Transactions are transparent due to

CHAPTER 5

Security Enhancement of Smart Grids using Blockchain Technology

Praveen Nelapati[1], Nitesh A. Funde[2,*], Anusha Viswanadapalli[3], Khushboo Jain[4] and **Meera Dhabu[5]**

[1] *School of Computer Science & Engg., VIT-AP University, Amaravati, Andhra Pradesh, India*

[2] *Department of AI, SardarVallabbhai National Institute of Technology (SVNIT), Surat, Gujarat, India*

[3] *Department of Computer Science & Engg., VFSTR University, Andhra Pradesh, India*

[4] *Department of Computer Science & Engg., Indian Institute of Information Technology, Nagpur, India*

[5] *Department of Computer Science & Engg., Visvesvaraya National Institute of Technology, Nagpur, India*

Abstract: The smart grid idea is implemented as a modern interpretation of the traditional power grid; it enables a two-way flow of electricity and data, including energy management, using different entities such as smart meters, appliances, and renewable energy resources. The critical nature of smart grids evokes traditional network attacks. Physical attacks, cyber-attacks, and natural disasters are significant threats to smart grid deployment. These threats can lead to infrastructural failures and various problems, including blackouts, energy theft, breaches of customer privacy, and endangering safety of operating personnel.

The blockchain has some significant features, making it an applicable technology for smart grid standards to solve security issues and trust challenges. We categorize the blockchain applications in the smart grid into three categories: energy trading, infrastructure management, and smart-grid operations management. We present different methods for security enhancement of Smart Grids (SG) using blockchain technology. A wide range of energy applications have suggested a suitable blockchain architecture in smart grid operations, a sample block structure and the potential blockchain technicalities employed in it. There is a need to critically examine the security issues aimed at preventing possible threats or failures. Various security challenges and threats are discussed with respect to their possible sources of occurrence. The important research problems and possible future research directions are presented for addressing smart grid security concerns using blockchain.

[*] **Corresponding author Nitesh A. Funde:** Department of AI, SardarVallabbhai National Institute of Technology (SVNIT), Surat, Gujarat, India; E-mail: nitesh.funde@aid.svnit.ac.in

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

Keywords: Blockchain, Consensus, Smart grid, Security, Threats, Trust.

INTRODUCTION

Facing significant challenges in the last few decades, conventional energy systems, which depend on fossil fuels, have encountered issues such as carbon emissions, long-distance transmission, pollution, and the imminent risk of an energy crisis [1]. With the ongoing development of human productivity, the power system has emerged as an indispensable infrastructure for human society, and the need for power resources is steadily increasing. The traditional power grid relying on coal, oil, and other fossil energy has served mankind for several years. However, the utilization of fossil energy has the problem of great harm to the environment, and the traditional power grid has the defects of poor interactivity, low degree of intelligence and security [2]. These defects often lead to huge economic and property losses for human society. In the pursuit of constructing a sustainable society to overcome these challenges, two primary potential solutions emerge: first, the integration of renewable energy resources and the enhancement of energy usage efficiency.

With the development of modern information, communication, computer and control technology, the application fields of these technologies have also been continuously expanded. The application of advanced information and communication technologies to power systems has been recognized by various countries, resulting in qualitative changes in power systems. New power grid architecture, namely smart grid [3] Smart grid, also known as knowledge-based grid, modern power grid, “Grid 2.0”, the definition given by the IEEE of the Chinese Academy of Sciences is: smart grid includes various equipment for power generation, transmission and distribution network, and energy storage. It merges seamlessly with the existing power grid through the integration of cutting-edge sensing and measurement technology, computing technology, network technology, communication technology, and automation and intelligent control technology. It is capable of monitoring and controlling the state of all equipment. It improves reliability, sustainability, and efficiency of power generation, transmission, and distribution [4].

Energy internet in smart grids helps enhance access to distributed and large-scale energy resources. It includes methodologies that can enhance the efficiency, utilization, and reliability of the energy system [5]. The Energy Internet addresses energy-related challenges by incorporating the Internet of Things, ICT, components of power systems, and other power networks [6]. In the evolution of the smart grid within the development of the Energy Internet System (EIS), decentralization is a crucial fundamental requirement in alignment with its vision

[7]. Nevertheless, the decentralized system, characterized by a multitude of components and intricate connections, poses security, privacy, and trust challenges, necessitating the exploration of new and innovative technologies for resolution.

In contrast, blockchain, being a transpiring and promising technology, opens up new prospects for the development of decentralized systems [8]. Blockchain technology, characterized by its decentralization, operates without the necessity for a central trusted authority in its management. Instead, interactions among multiple entities within the network facilitate the creation, maintenance, and storage of a chain of blocks. Within a decentralized system, every entity can confirm the integrity of the chain's order and data, preventing tampering. This feature not only introduces redundancy but also enhances resilience to system failures and cyber-attacks, providing solutions to multiple challenges found in centralized systems [9]. While initially introduced and primarily associated with digital currencies, blockchain has gathered remarkable attention in several non-monetary applications because of its outstanding properties. Simultaneously, blockchain is propelling the development of a secure, trusted, privacy-preserving smart grid [10]. This chapter is organized as follows: It begins with an overview of the smart grid, followed by a section on blockchain technology. Next, we discuss the mechanisms of blockchain as they pertain to the smart grid. Finally, we detail the categories of blockchain, consensus mechanisms, and their benefits, concluding with a summary.

SMART GRID

The smart grid encompasses the entirety of the electricity transmission, generation, and distribution network system in a unified framework, making the entire system more intelligent, efficient, and secure. As the global demand for clean energy rises, ongoing enhancements to electricity distribution grids lead to the automation of smart grid systems. This automation extends to substations, where advanced switching functionality is employed to increase the voltage source and overall power of the grid. In substations, a phasor measurement unit is incorporated to make time- and location-specific measurements of transmission line voltage, current, and frequency. The synchrophasor measurement data is collected at a rate of 30 to 120 samples per second, offering a more efficient solution compared to the data recorded every two to four seconds in traditional power systems, ultimately improving the efficiency of the power grid.

The increasing worries about emissions of gas emissions, such as carbon dioxide (CO_2), and the need for more reliable and efficient transmission of power and distribution system are propelling the development of a national power grid

CHAPTER 6

AI-enabled Security in Blockchain System

Pawan Whig[1,*], Ramya Thatikonda[2], Jhansi Bharathi Madavarapu[2] and **Ashima Bhatnagar Bhatia[1]**

[1] *Vivekananda Institute of Professional Studies-TC, New Delhi, India*

[2] *University of the Cumberlands, Williamsburg, KY 40769, USA*

Abstract: Blockchain allows the safe interchange of sensitive data without the need for duplication, reducing inaccuracies in medical records and gaining time through time savings. The data is timestamped as well, further enhancing its security. Blockchain technology is able to increase the safety and effectiveness of payments in a broad range of medical contexts. As a consequence, only those who have given permission to read or change health data are able to do so. In a safe and secure health sector with Automation tech, this book chapter suggests using chain technologies to construct an information and knowledge repository and retrieval mechanism for patients and healthcare providers. For the benefit of scholars within the same area, a case study focused on security and privacy concerns of the virtual world will be explored in this book chapter.

Keywords: AI, Accuracy, Blockchain, Privacy, Security, Technology.

INTRODUCTION

Blockchain is a decentralized digital ledger that records transactions across a network of computers, as demonstrated in Fig. (**1**). Individually, a chunk in the chain covers an amount of dealings, and every block is linked to the previous one through the use of cryptography [1]. This creates a secure and tamper-proof record of all transactions, as any alteration to a block would require the alteration of all subsequent blocks, which is computationally infeasible. The most well-known application of blockchain technology is Bitcoin, but it has many other potential uses, such as supply chain management and voting systems [2].

Blockchain technology was first proposed in 2008 by an individual or group of individuals using the pseudonym Satoshi Nakamoto. In a whitepaper titled "Bitcoin: A Peer-to-Peer Electronic Cash System," Nakamoto outlined a new electronic payment system that would use a decentralized, digital ledger to record

* **Corresponding author Pawan Whig:** Vivekananda Institute of Professional Studies-TC, New Delhi, India; Tel: +91-9811908699; E-mail: pawanwhig@gmail.com

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

transactions. This ledger, known as the blockchain, would be maintained by a network of computers rather than a central authority [3].

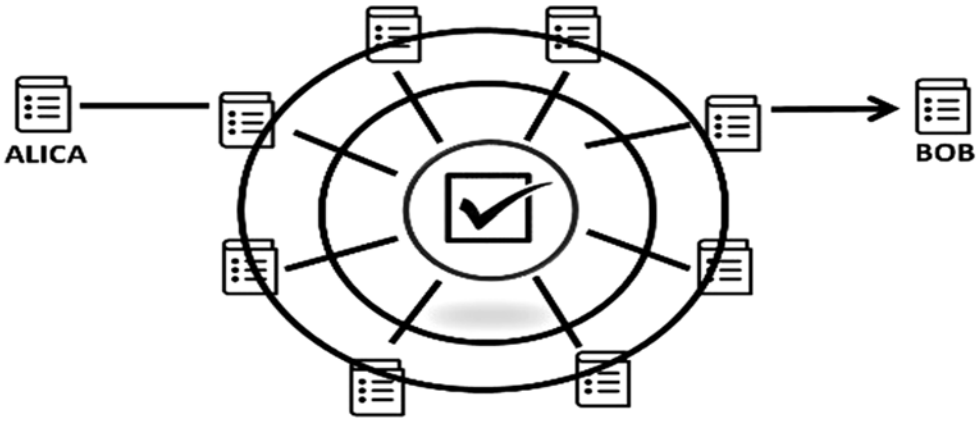

Fig. (1). Decentralized digital ledger.

The first blockchain-based cryptocurrency, Bitcoin, was launched in January 2009. The success of Bitcoin led to the development of other blockchain-based cryptocurrencies, such as Ethereum and Litecoin.

Over the next few years, the potential uses of blockchain technology were explored in various industries, including finance, supply chain management, and voting systems. In 2015, the Linux Foundation created the Hyperledger project, an open-source collaboration aimed at advancing the development of blockchain technology for use in business [4].

In recent years, blockchain technology has gained increasing attention and investment from major companies and organizations, as demonstrated in (Fig. (**2**). The use cases for blockchain technology are expanding, for example, in the field of digital identity, digital asset management, decentralized finance, and many more.

Blockchain technology is used for a variety of reasons, including:

Decentralization

Blockchain technology is decentralized, meaning it is not controlled by any single entity, as illustrated in Fig. (**3**). This decentralization enhances its security and resistance to tampering or hacking.

Fig. (2). Blockchain's importance to business.

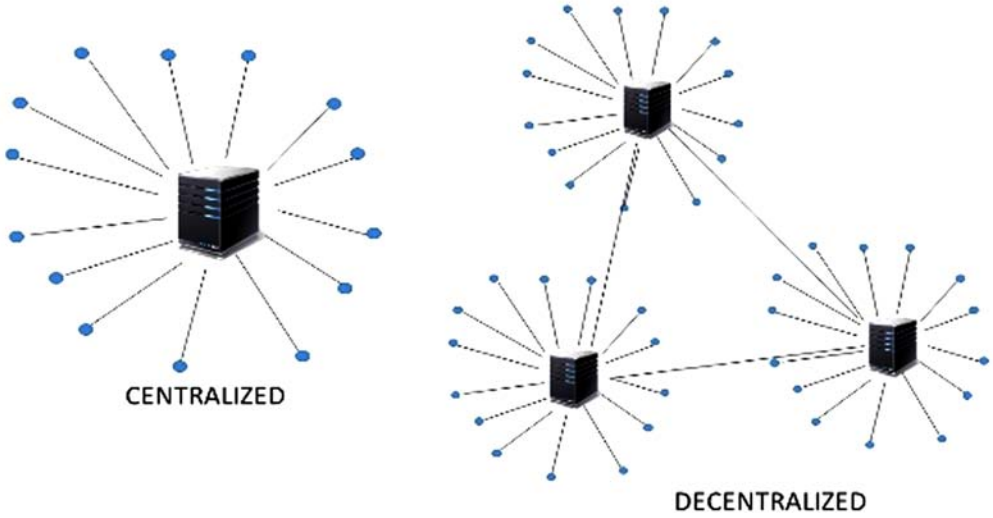

Fig. (3). Centralized *vs* Decentralized.

Decentralization is one of the key features of blockchain technology. In a decentralized system, power and control is distributed among multiple parties rather than existence centralized in a single entity [5]. In the case of blockchain technology, the decentralized countryside of the network means that transactions

CHAPTER 7

Cyber Attacks on Big Data System

Lipsa Das[1,*], Suman Avdhesh Yadav[1], Deepshikha Bhargava[1] and **Khushi Dadhich[1]**

[1] *Amity University, Greater Noida, UP, India*

Abstract: The way in which companies process and leverage large swaths of information has completely changed due to Big Data technologies. Furthermore, these technologies have also created new openings for global hackers who aim to compromise the hefty vaults of information. The sheer scale and complexity of Big Data systems and their frameworks have made them difficult to assault, which has also opened them up to a variety of attacks. For instance, cybercriminals deploy various forms of malware, such as viruses and worms, that erode the integrity of the information. Alternatively, perpetrators use aggressive approaches, sometimes relying on victims to go outside their networks and expose their critical information. Their vulnerability often increases due to an absence of powerful encryption, which can further increase their susceptibility by focusing on the complexity of these systems. The effects of such attacks are significant and extend far beyond merely operational continuity. Even greater problems relate to an organization's confidentiality and the integrity of its data. However, by focusing on advanced threat detection methods, stronger access controls, and encrypted Big Data systems, organizations should be at least partially able to reduce these threats. Moreover, some semblance of the current security may be maintained throughout the Big Data systems' lifetimes. As society increasingly relies on large-scale data processing, it becomes essential to overcome the unique obstacles presented by cyber threats. Doing so ensures that these critical systems can continue to operate efficiently and reliably.

Keywords: Bigdata, Cyberattack, Security.

INTRODUCTION

Smartphones, laptops, connected watches, the Internet, social networks, connected devices, and streaming technologies have all been adopted, and as a result, daily life has become increasingly digital and is producing an enormous amount of digital content that is constantly expanding. Since data has developed significantly during recent years and is now being demanded to handle terabytes, petabytes, and now zettabytes, such new approaches are needed to manage such a large qua-

[*] **Corresponding author Lipsa Das:** Amity University, Greater Noida, UP, India; E-mail: lipsaentc9@gmail.com

Sonali Vyas, Shaurya Gupta, Vinod Kumar Shukla & Chinwe Peace Igri (Eds.)

ntity of data. A recent technique of information management, known as Big Data, has begun because of this massive data output. Big Data is the term for the content that is used by many businesses across various industries that want to quickly and automatically extract strategic information [1].

Since the big data environment is so vast, intricate, chaotic, full of inaccurate and noisy information, and heterogeneous, it may alter how standard statistical and data analysis methods are used. The truth is that having more data does not always equate to having more valuable data [2]. Nevertheless, it appears that employing big data makes it possible to acquire more information in order to extract some valuable data. Any storage with a significantly high velocity, volume, correct, unique, and meaningful information is referred to as "big data" in general. Examples include viewing records, caches, networking site information, webcam logs, media viewed, and collected sensitive data. Even in real-time, at the appropriate moment, businesses may be able to obtain invaluable information from this type of storage. Top firms place greater emphasis on achieving their objectives since doing so enables governments to anticipate unforeseen threat threats and alter the portfolio to satisfy client wants. However, organized databases and conventional methods cannot be used to analyze and visualize such a large volume of huge datasets since they require enormous simultaneous computing methods. Consequently, it became important to offer new methods and tools in line with such an expansion. Big Data analytics refers to these methods and tools.

Applying cutting-edge parallel and analytical techniques to manage exceedingly large and different records with a variety of contents is known as big data analytics. By processing any large amount of combined unorganized, semi-organized, and organized information that is constantly evolving and challenging to manage using standard database techniques, big data analytics solutions enable generating considerable advantages and priceless discoveries. Big Data offers a wide range of businesses and decision-makers tons of benefits, but it also raises serious safety and privacy risks. Modern analytics solutions must collect, preserve, analyze, display, and share diverse data from all practical and available sources, which presents a difficulty. The end effect is that Internet users are substantially more exposed as a result of the aggregation and exploration of specific behavioral data. In other words, it is possible to collect and extract considerably more data from these technologies than is necessary, which causes several privacy and security issues [3].

The security and privacy of these records are of utmost importance and are top priority. To address these issues, the research community must propose and put into place robust protection methods that allow for the use of big data without

compromising privacy and safety. Although there have already been some initial attempts to provide a secure layer when working with big data, more potential solutions should be investigated to cover all bases and create a solid big data platform [4].

Big data-related cybercrime is growing at an unprecedented rate, which has a negative impact on the Internet sector and global data. The increasing use of offensive and assault strategies by cyber criminals, as well as the increasing importance of data- and intelligence-driven adversaries, show that conventional countermeasures against cyber threats are losing their efficacy. Big Data is widely used for marketing, fraud, and criminal activity detection, epidemic intelligence, *etc*. In the big data context, cyber-attacks such as malware and injection are increasingly widespread. Due to the emergence of malware, social engineering, Advanced Persistent Threats (APTs), and other forms of digital crime, attacks are getting increasingly complicated [5]. Since data from online social media are easily accessible, there is always potential for gain. Social networking sites are commonly used to collect individual information and earn immoral profits. Cyberbullying, phishing, and identity theft are all prominent ways that social networking sites are attacked. As threat actors multiply over time, modern systems are increasingly vulnerable to both state-sponsored initiatives and hordes of attackers. Due to the inability of current security technologies to identify them, the threat of previously undiscovered cyberattacks has recently increased. Cyberattacks in the past had the straightforward aim of releasing personal data by targeting the personal computer or crashing the system. However, current hacking attempts have shifted from targeting large-scale systems like vital infrastructure and government organizations to attacking information leaks and service destruction. In other terms, the solutions used today to defend against these assaults are founded on quite simple pattern recognition methods. This results in a relatively low detection rate and a rise in false negatives in the case of novel and previously unidentified attacks. Organizations are progressively researching cutting-edge cybersecurity tactics to counter these increasingly complex assaults. Instead of merely depending on sequentially scanning potential attack paths, production is working to set up an organization that provides constant tracking and extraction from their architecture. In order to provide pertinent knowledge which not only permits risk recognition and reaction but also forecasts threats prior to actual networks being damaged, this data may be fed into real-time behavioral analytics and machine learning technologies.

This article examined several large data assaults, storage mechanisms, and real-time analytics techniques.

SUBJECT INDEX

C

D

E

F

G

H

I

K

L

M

N

O

P

R

S

T

V

W

Sonali Vyas

Dr. Sonali Vyas, an Associate Professor in the School of Computer Science at the University of Petroleum and Energy Studies, Uttarakhand, has over 14 years of experience in Computer Science. Her expertise includes Healthcare Informatics, Data Science, Database Virtualization, Data Mining, and Data Analytics. A Senior Member of IEEE and author of Smart Health Systems (Springer), Dr. Vyas has over 120 publications and has edited several books in emerging technologies. She holds four patents, frequently speaks at national and international conferences, and is actively involved in academic, research, and editorial roles in her field.

Vinod Kumar Shukla

Dr. Vinod Kumar Shukla, an Associate Professor and Head of Academics at Amity University, Dubai, specializes in Semantic Web and Ontology. Ranked in the top 2% of scientists in the Stanford list 2022, he has numerous publications, international patents, and books in Industry 4.0 and IT. He has conducted training for various organizations and completed the General Management Programme at IIM Ahmedabad. Dr. Shukla is also a member of IEEE and a contributor to academic and industrial collaborations.

Shaurya Gupta

Dr. Shaurya Gupta, an Assistant Professor at the University of Petroleum and Energy Studies, Dehradun, has over 11 years of experience in teaching, administration, and research. He earned his Ph.D. in IT from Amity University, Rajasthan, and specializes in Delay Tolerant Networks, IoT, Machine Learning, and Ad Hoc Networks. Dr. Gupta has published extensively in reputed journals and conferences and is actively involved in preparing academic materials for UG and PG students.

Chinwe Peace Igiri

Chinwe Peace Igiri, a lecturer at Mountain Top University, Nigeria, and guest lecturer at Cavendish University Uganda, has 13+ years of experience in teaching, research, and administration. Holding a Ph.D. in Software Engineering from Amity University, she specializes in Computational Intelligence, Blockchain, Data Science, and AI. She has published 15 peer-reviewed articles, holds professional certifications in blockchain technology, and is a member of COREN, IEEE, and ACM.