A systematic literature review: Cybersecurity through Blockchain

***Abstract*—Blockchain Technology**
***Index Terms*—Blockchain, cybersecurity**

I. Introduction

Blockchain Technology has its origins in the paper "How to Time-Stamp a Digital Document" published in 1991 by Stuart Haber and W. Scott Stornetta [1]. But it came into limelight after the release of the whitepaper "Bitcoin: A Peer-to-Peer Electronic Cash System" by the pseudonymous author Satoshi Nakamoto in 2008 [2] followed by the release of its implementation in 2009. Bitcoin gained popularity in no time as increasing number of people were willing to invest in bitcoin. It was the first cryptocurrency that allowed reliable financial transactions without the need of a trusted central authority, such as banks and financial institutions [3]. Blockchain is the technology underlying the Bitcoin. Simply put, a blockchain is a cryptographically linked chain of blocks, each of which contains time-stamped records. It is a decentralized system where each user has the access to all records. The records stored on a blockchain are immutable and are verified by multiple parties through various consensus protocols. A blockchain has been described as a value-exchange protocol [23]. It has the potential to address the pressing issues in the exchange of digital assets such as Lack of Trust, Double-spending, Repudiation and Theft (including fraud). Blockchain Technology has the potential to change several other industries apart from the Finance industry. One of the fields that naturally pairs with blockchain is Information Technology. Almost every industry today - Healthcare, Pharmaceuticals, Agriculture, Automotive, Media, Finance, Manufacturing, Transportation, Government, Trade, Education has embraced Digitization and is dependent on the Information Technology industry. Digitization has revolutionized the way we transact, communicate and commute making it faster and more efficient. Digitization makes an industry more productive, but it also introduces risks and vulnerabilities. The IT risks introduced by Digitization include hardware and software failure, human error, spam, viruses and malicious attacks, as well as natural disasters such as fires, cyclones or floods [24]. Such events have a negative impact on the company and affect its clients. Many solutions have been proposed to address various IT risks till date. But the risks still persist, a major reason being that many solutions rely on centralization. Blockchain has a decentralized structure. It is a type of distributed ledger. Each block is linked to the previous block by a cryptographic hash. Any tampering with one block causes all the further blocks in the chain to become invalid. This rigid infrastructure of blockchain has the potential to offer cybersecurity solutions to the threats faced by various components of Information Technology such as IoT devices, networks and data storage and transmission.

A. Prior research

B. Research goals

C. Contributions and layout

II. Research Methodology

III. Risks in Information Technology

In the IT field, the most valuable asset is considered to be an individual's data. This is a data-driven world. Companies request for the data of a customer which can help to enhance the services provided by the company. The users have no choice but to provide their data and hope that it is not being misused or shared with other parties. Some of the main risks in the IT field are:

A. Data Security and Integrity

Data Security risk is one of the most crucial risks in the IT field. Over the years, the number of data breaches and their magnitudes have increased exponentially. They are associated with the theft of medical information, account credentials, confidential emails, and other forms of sensitive information. A consumer has to face several consequences due to a data breach such as identity theft, fraudulent credit card activities and temporary account cancellations. One in three data breach victims later go on to experience an identity crime [25]. The average data breach now costs a company up to $ 3.82 million [26]. The aftermath of a data breach at a major organization is chaotic, with a fall in share prices of the company, underperformance for nearly a year and damage to the trust of its customers. 2019 witnessed some of the largest data breaches of the decade, including the data breaches at Equifax, Facebook, Instagram and many Healthcare organizations. In addition, there are several data breaches that go unnoticed by companies and uninformed to the public. Thus, there is a large room for improvement in the techniques currently being implemented to safeguard data.

B. User Data Privacy

Technology is evolving rapidly. Increasing amounts of data is being collected for enhancing user experience, but this has given rise to potentially serious privacy concerns. It is predicted that more than half of global enterprises currently using some form of cloud solutions will have adopted a full cloud strategy by 2021 [27]. This will lead to increased

dependence on third-party providers for data storage and data privacy. Many businesses are embracing Artificial Intelligence to drive value through automation due to which there is collection of more data and hence increased privacy failures that arise with the creation of new types of metadata. Another area of concern is the weak security configurations of Internet-of-Things (IoT) devices. It is predicted that by 2025 the total installed base of IoT connected devices will be close to 75 billion. IoT technology increases the number of access points multiplying the risks of compromise of personal information.

C. Cloud computing risks

It is predicted that by 2025 the amount of data worldwide will reach an overwhelming 175 zettabytes, almost 49% of which will be stored in public cloud environments. But cloud computing is faced with several serious issues of security, compliance, governance, migration, licensing and multiple cloud management. Examples of cloud data breaches include the 2018 Aadhar India National ID database breach which exposed the data of 1.1 billion citizens and the Cambridge Analytica accessing Facebook data to profile voters. There is heavy centralization in the cloud environment due to the dominance of the giants like Google Cloud, Amazon Web Services (AWS), Alibaba, IBM and others. This can have a huge impact as demonstrated by the 2018 cloud outage incidents of Amazon Web Services (AWS), Google Cloud and Microsoft Azure. When an organization uses external cloud services, the data visibility and control gets reduced. Cloud Service Provider (CSP) APIs are accessible via Internet making them more vulnerable to attacks that lead to compromise of cloud assets. Other threats to cloud computing include failure of separation among multiple tenants, incomplete data deletion, infeasible migration from one CSP to another, unauthorized data access to insiders and accidental data deletion by the CSP. It is important to address these cloud computing risks to protect the enormous amount of data that we store on the cloud daily.

D. Internet-of-Things security

IoT devices collect user data to personalize user experience and provide utility to the users. The dark side of this increasing data collection is the possible creation of a virtual biography of user activities, exposing life style patterns and private information [4]. Several intrinsic features of IoT amplify its security and privacy challenges including: lack of central control, heterogeneity in device resources, multiple attack surfaces, context-aware and situational nature of risks, and scale [4]. It is estimated that the total number of IoT connected devices in the world will be around 38.5 billion by the end of 2020 [28]. With the increase in number of IoT devices, there is an increase in the number of cyber-attacks and botnets. IoT devices have a weak security configuration due to which they fall victims to malware which allows the attacker to control the devices forming an IoT botnet. Attackers make use of such botnets to launch Distributed-Denial-of-Service (DDoS) attacks. Linux.Aidra, Bashlite, Mirai and IRCTelnet are some of the largest IoT botnets till date. The Mirai Botnet was responsible for DDoS attack on Dyn Server in 2016, it was the largest of its kind causing disruption of the internet service across US and Europe. Several solutions have been proposed but still security of IoT devices continues to be a pressing issue that needs to be addressed urgently.

E. Trust

The internet was not built with trust-building in mind [29]. The two major players that influence online trust are Corporations and Attackers. Corporations are concerned about their users' data security but their primary interests are not the same as those of their users. The attackers constantly invent new ways to break into a system and access unauthorized data. As the users become aware of the growing number of data breaches taking place, they become distrustful towards online services. However, most users continue to do online interactions as they make life easier. There is a belief among users that no one wants their data, they trust that the corporations will ensure data security and hence give in to risks. Lack of trust is a major issue in centralized systems. For example, in the Public Key Infrastructure (PKI), both the transacting parties must trust the Certificate Authority (CA) to sanction valid certificates for the exchange of digital assets in a secure way. When we make online payments, we need a trusted-third-party like a financial institution or central bank to complete the transaction. The trust between users and service providers will soon diminish as the number of data breaches, identity thefts and hacking incidents grow.

F. Single Point of Failure (SPoF)

A single point of failure (SPOF) is a part of a system that, if it fails, will stop the entire system from working (Wikipedia). Single Point of Failure is synonymous to Centralization. It threatens the availability and reliability of centralized systems. An example of low efficiency and risks introduced due to centralization is the antivirus agencies of today, which follow a centralized model. A small team is responsible for detecting all threats and malwares for their users. With the increasing number of users, this model puts a large burden on the team thus decreasing the efficiency of detecting threats and makes the user device vulnerable to malware. SPOF maybe be caused due to hardware, Internet Service Provider (ISP), an offsite data storage location like the cloud. People also constitute to SPOFs in many organizations. SPOFs provide the attackers an easy way in to the database of an organization. Another possibility is that an SPOF may get triggered accidentally due to human errors. In either case, there is a breakdown of the system causing the company to incur heavy loss in terms of capital as well as trust of its consumers.

IV. Blockchain in IT risk management

A. Data Security

The article [19] explains the innovative blockchain solutions being implemented in Estonia to enhance data security. Keyless Signature Infrastructure (KSI) is assumed to be used

in Estonia to store public data to ensure data security. [20] proposes Data Anchoring to overcome the inefficiency associated with storing large files and databases on the blockchain. Rather than storing a full digital asset on the chain, a unique hash for that digital asset is stored on the chain, along with some method of identifying the asset you are trying to protect. This method does not prevent data-tampering but it provides a powerful tamper-detection as well as asset authentication mechanism. Data anchoring along with frequent recovery points via snapshots, mirrors, backups archives increases the levels of system integrity in an IT environment. The study [7] intends to identify the most significant, ingrained risks posed to transactions in a digital environment and to understand how some specific characteristics of blockchain technology could potentially address those risks. Currently, the major risks with the exchange of digital assets are identified to be Lack of Trust, Double-spending, Repudiation and Theft (including fraud). The research focuses on how the characteristics of blockchain technology such as Peer-to-Peer (P2P) network, Distributed ledgers, Consensus mechanism, Asymmetric cryptography, Immutability could address the identified risks. The author has put together the risks and the corresponding blockchain characteristics in a mapping which serves as a quick reference matrix for potential users. The researchers in [8] focus on the advantages and limitations of blockchain-based reputation systems to prevent rating fraud. Blockchain technology is effective in protecting customers' privacy, thus encouraging them to give honest feedback. Blockchain systems are effective against rating fraud in non-computational and content-driven reputation systems as verification of claims is possible. Blockchain-based reputation systems can be used to prevent "bad mouthing", "whitewashing" and "sybil" attacks, but are not resistant to "ballot stuffing", "constant" and "camouflage" attacks.

B. Cloud computing

In [21], the author has emphasized the major role that clouds will play in data storage by 2025. The article suggests use of Blockchain technology to enhance cloud security and prevent inflation of costs of services. The giants like Google Cloud, AWS, Alibaba, IBM and more hold majority of cloud computing making it heavily centralized. Adopting blockchain technology transfers the data back to the users even though there is centralization due to cloud computing service providers. Ankr, Transcodium, Dfinity, Oasis Labs, BonusCloud and Sia are the recent projects that combine blockchain and cloud computing. The world's largest computing company, Avogadro Corporation has embraced blockchain-based cloud computing platform.

In [9] the researchers have proposed a blockchain based cloud computing model to immediately detect the corruption of data. Data is stored on the cloud and its encrypted form is stored on the blockchain. A blockchain function gets invoked if data is tampered with and the user gets informed automatically. This method is more efficient in the computational cost and overhead as compared to the methods proposed by by Atneise et. al [5] and Erway et. al [6]. It provides a guarantee of corruption detection as well as a lower detection time.

C. IoT

The blend of Internet-of-Things and decentralized architecture like Blockchain technology to tackle IoT security risks is catching the attention of many researchers. Based on IoT functionalities, it runs on three architecture layers which are perception, network, and application layers [10]. The application layer is exposed to the outside network and is vulnerable to attacks like Denial of Service (DoS) attack, privacy leak, malicious code, and social engineering. To secure the application layer, researchers in [10] have proposed "Dynamic Self-Protection" model which integrates Blockchain technology and adaptive security mechanism features. The features of blockchain such as decentralization, anonymity and security have the potential to address privacy and security issues in IoT. But blockchains require high computation power and introduce significant overhead traffic which is problematic for IoT devices. Hence, authors in [11] have proposed a blockchain-based architecture for IoT that delivers lightweight and decentralized security and privacy. They demonstrate the working of the model via a smart home application case study. The data storage, access and modifications are based on the interactions between the smart home, the overlay network and the cloud storage which form the three tiers in the architecture. The researchers in [12] have proposed DroneChain, an architecture that integrates blockchain with drone-based IoT applications to ensure secure drone communication and provide data assurance, resilience and accountability. The experimental results confirm that DroneChain has a linearly increasing response time with varying number of drones as well as data size providing better scalability. DroneChain also has an acceptable average response latency. Li et al. [13] proposed a low-cost identity authentication and security mechanism for IoT devices based on blockchain technology. They verified the model by implementing a prototype system based on Hyperledger Fabric. The researchers in [14] have proposed the NeuroMesh solution. It combines neural and mesh networks along with the Bitcoin blockchain and machine learning algorithms to secure the IoT devices from Botnet attacks by training the IoT devices with Mirai botnet – essentially working like a vaccine for IoT bots. The NeuroMesh makes use of Blockchain technology to ensure safe communication between Command and Control centre and IoT devices, as well as to store information about blacklisted and whitelisted IP addresses and malware signatures. It is also used to broadcast information about new threats to all devices providing a global threat sharing and intelligence mechanism for all devices irrespective of their owners. [22] proposes the use of Blockchain Technology to prevent Distributed-Denial-of-Service (DDoS) attacks. The centralized client/server model of the Internet causes the entire system to fail due to the failure of a single server thus facilitating DDoS attacks. A Blockchain solution will replace login credentials with public key cryptography making it difficult to scan and compromise devices. User's

private key will become necessary to communicate with other devices. By limiting the authority to install firmware to the manufacturer using his private key, the chances of installation of malware would become almost none. Storing public keys on the Blockchain will help IoT devices to authenticate login requests. Blockchain can thus establish a secure P2P network wherein the attacker will not be able to publish DDoS attack launch instructions.

D. Trust, Centralization and Single Point of Failure issues

The current Public Key Infrastructure (PKI) authentication is mainly dependent on Certificate Authorities (CA) which have to be trusted by domain owners as well as operators. There is a disparity of rights between users and CAs. Methods such as Domain Name System Security Extensions (DNSSEC) and Certificate Authority Authorizations (CAA) are trust-based and are prone to single point of failure. The research [15] proposes a domain authentication system based on blockchain technology known as AuthLedger. The AuthLedger consists of five entities – CA, Domain Name Server (DNS), Browser Extention, Validating Authority, Blockchain. The domain authentication procedures are based on time and count. Rules have been formulated to prevent misbehavior of validator nodes as well as clients which also include incentives to promote node honesty and attract more nodes to join the system. The research describes the implementation prototype with the help of Ethereum smart contracts and Solidity language. The researchers in [16] have proposed TrustChain – a privacy preserving blockchain. Transactions in TrustChain are based on the measure of the trustworthiness of a user through three Trust Metrics – knowledge, experience and reputation. The TrustChain is designed to work in compliance with the GDPR legislature to ensure privacy of user's data through the implementation of ZKP, data encryption and user anonymity. The consensus protocol involves "Trust Bloggers" who verify the transactions with a method similar to BFT but without a Central Authority. The TrustChain eliminates the requirement of high computational power making it suitable for IoT networks.

E. Risk Control

In [17], the researchers have proposed a blockchain-based Risk and Information System Control (RISC) framework as a mechanism for sharing risk information among insiders, IoT devices and information systems. The research has designed three types of risk smart ledgers and the Merkel tree is used to establish relationship between them. To avoid continuous monitoring, the study designs three risk smart contracts for automatic risk calculation and approval flow control during each of the four stages of the life cycle. Blockchain technology effectively handles Backtracking, Tracking, Falsification and Multi-trust issues in the traditional RISC making RISC more efficient. The research [18] proposes a blockchain-based security framework – Sapiens Chain. In combination with artificial intelligence, this framework makes it possible to detect vulnerabilities automatically and handle website, application and blockchain securities simultaneously. In the proposed framework, the users submit tasks through the browser, fog nodes distinguish the tasks and assign fragments of tasks to selected nodes. Finally, results are gathered into a report. The experimental data confirms that Sapiens Chain framework can distinguish between websites at high and low risk levels as well as it can detect vulnerabilities and provide corresponding suggestions. The study discusses the applications of the proposed framework such as website, application and smart contract securities and Shared Economy.

V. Analysis of the findings

VI. Potential benefits

VII. Challenges and open issues

VIII. Conclusion

References

[1] S. Haber and W. S. Stornetta, "How to time-stamp a digital document," *Journal of Cryptology*, vol. 3, pp. 99–111, 1991.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Cryptography Mailing list at https://metzdowd.com*, 03 2009.

[3] F. Tschorsch and B. Scheuermann, "Bitcoin and beyond: A technical survey on decentralized digital currencies," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 2084–2123, thirdquarter 2016.

[4] A. Dorri, S. Kanhere, and R. Jurdak, "Blockchain in internet of things: Challenges and solutions," *arxiv*, 08 2016.

[5] G. Ateniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. Peterson, and D. Song, "Remote data checking using provable data possession," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, Jun. 2011. [Online]. Available: https://doi.org/10.1145/1952982.1952994

[6] C. C. Erway, A. Küpçü, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession," *ACM Trans. Inf. Syst. Secur.*, vol. 17, no. 4, Apr. 2015. [Online]. Available: https://doi.org/10.1145/2699909

Research Papers

[7] T. Mari, "Blockchain technology: Addressing the risks of digital assets exchange," *Thesis (MAcc)–Stellenbosch University*, 2018.

[8] Z. D. Cai, Y., "Fraud detections for online businesses: a perspective from blockchain technology." *Financ Innov 2, 20*, pp. https://doi.org/10.1186/s40 854–016–0039–4, 2016.

[9] E. A. Kanimozhi, M. Suguna, and S. Mercy Shalini, "Immediate detection of data corruption by integrating blockchain in cloud computing," in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, March 2019, pp. 1–4.

[10] B. H. AlDoaies and D. H. Almagwashi, "Exploitation of the promising technology: Using blockchain to enhance the security of iot." *2018 21st Saudi Computer Society National Computer Conference (NCC)*, pp. 1–6, 2018.

[11] A. Dorri, S. S. Kanhere, and R. Jurdak, "Blockchain in internet of things: Challenges and solutions," *ArXiv*, vol. abs/1608.05187, 2016.

[12] X. Liang, J. Zhao, S. Shetty, and D. Li, "Towards data assurance and resilience in iot using blockchain," *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, pp. 261–266, 2017.

[13] D. Li, W. Peng, W. Deng, and F. Gai, "A blockchain-based authentication and security mechanism for iot," *2018 27th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–6, 2018.

[14] G. Falco, C. Li, P. Fedorov, C. Caldera, R. Arora, and K. Jackson, "Neuromesh: Iot security enabled by a blockchain powered botnet vaccine," in *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, ser. COINS '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–6. [Online]. Available: https://doi.org/10.1145/3312614.3312615

[15] Z. Guan, A. Garba, A. Li, Z. Chen, and N. Kaaniche, "Authledger: A novel blockchain-based domain name authentication scheme," in *ICISSP*, 2019.

[16] U. Jayasinghe, G. M. Lee, MacDermott, and W. S. Rhee, "Trustchain: A privacy preserving blockchain with edge computing," *Wireless Communications and Mobile Computing*, vol. 2019, pp. 1–17, 07 2019.

[17] S. Ma, W. Hao, H. Dai, S. Cheng, R. Yi, and T. Wang, "A blockchain-based risk and information system control framework," in *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, Aug 2018, pp. 106–113.
[18] Y. Han, Z. Wang, Q. Ruan, and B. Fang, "Sapiens chain: A blockchain-based cybersecurity framework," *ArXiv*, vol. abs/1811.10868, 2018.

Industrial Reports and Articles

[19] V. Kumar, "How blockchain can make data secure for companies," *yourstory.com*, 2018.
[20] chainkit, *Introduction to Data Anchoring*. [Online]. Available: https://chainkit.com/data-anchoring
[21] MarkNetwork, *WILL BLOCKCHAIN DISRUPT CLOUD 2.0?* [Online]. Available: https://www.mark-network.com/will-blockchain-disrupt-cloud-2-0/
[22] Deloitte, *Prevention of DDoS Attacks*. [Online]. Available: https://www2.deloitte.com/de/de/pages/technology-media-and-telecommunications/articles/cyber-security-prevention-of-ddos-attacks-with-blockchain-technology.html

Refernce Links

[23] *Blockchain*. [Online]. Available: https://en.wikipedia.org/wiki/Blockchain
[24] IBM, *Information Technology(IT) Risk Management*. [Online]. Available: https://www.business.qld.gov.au/running-business/protecting-business/risk-management/it-risk-management
[25] S. Turner, *2019 Data Breaches*. [Online]. Available: https://www.identityforce.com/blog/2019-data-breaches
[26] IBM, *Cost of a Data Breach study*. [Online]. Available: https://www.ibm.com/security/data-breach
[27] B. Lee, *10 Privacy Risks Every Company Should Prepare for in 2018*. [Online]. Available: https://www.corporatecomplianceinsights.com/10-privacy-risks-every-company-should-prepare-for-2018/
[28] J. Research, *'INTERNET OF THINGS' CONNECTED DEVICES TO ALMOST TRIPLE TO OVER 38 BILLION UNITS BY 2020*. [Online]. Available: https://www.juniperresearch.com/press/press-releases/iot-connected-devices-to-triple-to-38-bn-by-2020
[29] L. Rainne and J. Anderson, *Trust will diminish because the internet is not secure and powerful forces threaten individuals' rights*.

To be completed.