Part I

Introduction to Blockchain

1

Introduction

Sachin S. Shetty,[1] Laurent Njilla,[2] and Charles A. Kamhoua[3]

[1] *Old Dominion University, Virginia Modeling, Analysis and Simulation Center, Norfolk, VA, USA*
[2] *US Air Force Research Lab, Cyber Assurance Branch, Rome, NY, USA*
[3] *US Army Research Laboratory, Network Security Branch, Adelphi, MD, USA*

1.1 Blockchain Overview

Blockchain technology has attracted tremendous interest from a wide range of stakeholders, which include finance, healthcare, utilities, real estate, and government agencies [1–5]. Examples of potential applications of this technology are claims processing, transparency and auditing of operations, identity management, supply chain provenance to address the threat of counterfeit products, and integrity of the information acquired from Internet of Things (IoT) devices. Blockchains are a shared, distributed, and fault-tolerant database that every participant in the network can share, but no entity can control. The technology is designed to operate in a highly contested environment against adversaries who are determined to compromise. Blockchains assume the presence of adversaries in the network and nullify the adversarial strategies by harnessing the computational capabilities of the honest nodes, and the information exchanged is resilient to manipulation and destruction. Blockchains facilitate the development of trustworthy networks in a trustless environment.

The premise of blockchain is that applications do not need a trusted central authority to operate and can function in a decentralized fashion. Blockchain enables exchange of information among distrusting entities. Blockchain enables trustless networks and allows entities to engage in transactions in the absence of mutual trust. There is an assumption that a communication medium could be compromised by insiders or outsiders. The reconciliation process between entities is sped up due to the absence of a trusted central authority or intermediary. Tampering of blockchains is extremely challenging due to the use of a cryptographic data structure and no reliability of secrets. Blockchain networks are fault tolerant, which allows nodes to eliminate compromised nodes.

Blockchain for Distributed Systems Security, First Edition. Edited by Sachin S. Shetty, Charles A. Kamhoua, and Laurent L. Njilla.

Blockchains have the following advantages over centralized databases: (i) ability to directly share a database across diverse boundaries of trust in situations where it is difficult to identify a trusted, centralized arbitrator to enforce constraints of proof of authorization and validity. In a blockchain, transactions leverage their own proof of validity and authorization based on a verification process managed by multiple validating nodes and a consensus mechanism that ensures synchronization; and (ii) ability to provide robustness in an economical fashion without the need for expensive infrastructure for replication and disaster recovery. Blockchain requires no configuration to connect and synchronize nodes in a peer-to-peer (p2p) fashion, with built-in redundancy and no need for close monitoring. It can tolerate multiple communication link failures, allows external users to transmit transactions to any node, and ensures disconnected nodes will be caught up on missed transactions.

Blockchain's distributed database maintains a continuously growing list of records, called blocks, secured from tampering and revision by distributed storage and continuous verification. The blocks contain a temporal listing of transactions that are stored in a public ledger using a persistent, immutable, and append-only data structure that is globally viewable by every participant in the underlying p2p network. When such an elegant data structure is considered to track data transactions in a distributed environment, the block structure contains attributes such as the set of user transactions, a timestamp, a reference to a previous block in the blockchain, Merkle root of the transactions, and so on. In this manner, the blocks are linked together to form a chain, where the hash of the previous blocks helps to maintain the integrity of the whole blockchain (Figure 1.1).

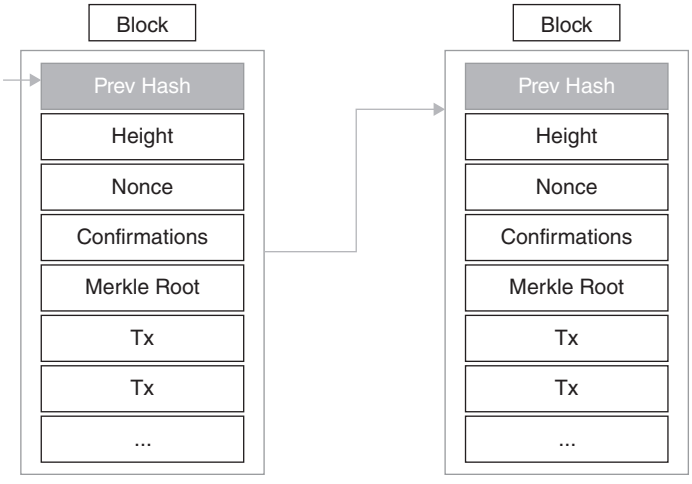

Figure 1.1 Block structure.

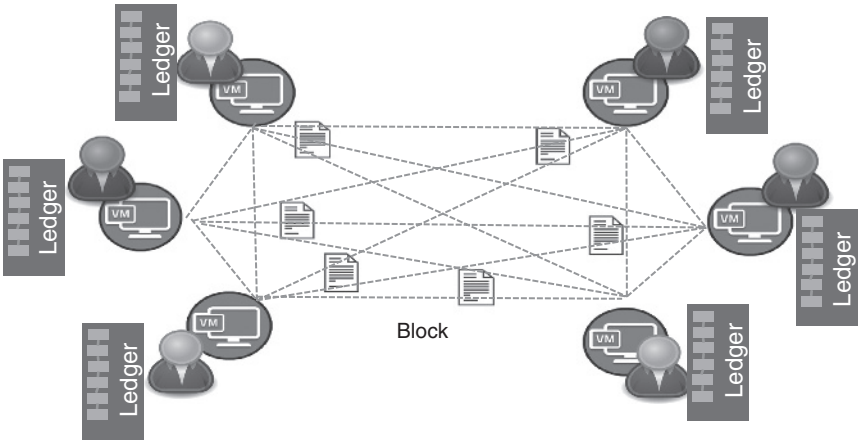

Figure 1.2 Blockchain architecture.

1.1.1 Blockchain Building Blocks

Blockchain technology's effectiveness hinges on the following three main components: a decentralized network, distributed consensus, and cryptographically secure algorithms. Figure 1.2 illustrates the basic blockchain architecture.

The key features of each of the components are as follows:

Decentralized Network—The function of the decentralized network is to ensure the propagation of transaction messages among the nodes responsible for maintaining the distributed ledger. The network protocol allows the transaction message to be broadcast from any node to all nodes in the decentralized network. However, the network is not a pure broadcast medium and allows nodes to propagate messages that represent valid transactions. The network can be part of a private or public blockchain that has ramifications on network performance and security. Irrespective of whether the blockchain is public or private, the decentralized network is based on a p2p architecture. The nodes can join and leave freely. There is no centralized arbitrator. The network has built-in redundancy and robustness to mitigate node and link failures.

Distributed Consensus—Blockchain uses consensus protocols over a decentralized p2p network for verification of transactions prior to adding blocks to the public ledger. The consensus protocol receives messages from the p2p network and inserts transactions in the distributed ledger. The consensus protocol is responsible for mining blocks and reaching consensus on their integration in the blockchain. The consensus protocol chooses the set of transactions that is accepted after passing a verification process. The verification process is determined by users and does not require a centralized

administrator. The consensus protocols ensure that the newly added transactions are not at odds with the confirmed transactions in the blockchain and maintain the correct chronological order. The newly added transactions that are waiting to be confirmed are packed in a block and submitted to the blockchain network for validation.

Cryptographically Secure Algorithm—The foundational component of blockchain technology is the cryptosystem. State-of-the art blockchains' cryptography systems use public key algorithms such as Elliptic Curve Cryptography, and message digests such as SHA3-256. In a typical blockchain application, an Elliptic Curve key pair that contains a public key and private key is generated based on Secp256K1 curves. The private key has the traditional usage of being kept secret and utilized to sign transactions. For instance, in the case of bitcoin use, when a user exchanges bitcoins with another user, the user will sign the transaction with their private key prior to announcing to the network. Once the transaction is signed, the miners in the network will use consensus algorithms to verify the validity of the transaction signature, and validation is achieved.

1.1.2 Blockchain Commercial Use Cases

Home IoT—"Smart home" is an emerging Internet of Things (IoT) application that aims to provide higher accessibility to all home accessories and personalized user experience for the appliances. To operate efficiently, IoT hubs in smart homes collect and analyze a lot of sensible data from the home area network of all smart devices. With the gathered data, it is easy to derive usage patterns and user behaviors in the home environment, thus creating a digital trail of families in smart homes. However, this information could easily fall into the wrong hands, or the vendors could use the information to promote additional products. Therefore, when more smart technologies are added to smart homes, there is an increased possibility of a severe privacy breach. Conventional approaches may fail to achieve credible security and privacy in IoT because the IoT framework has particular characteristics. These characteristics include decentralized topology, resource-constrained devices, limited network performance, and minimal security standards for IoT devices. The introduction of blockchain technology could potentially address smart home challenges in an efficient manner because of the following facts: blockchain does not rely on a centralized control; instead, it works in a distributed network setting that is similar to IoT. As a result, blockchain avoids the problem of single-point failure and improves scalability. Moreover, blockchain inherently offers the anonymity that is required in the IoT environment, where identities must be kept private. Irrespective of the several advantages produced by blockchain technology, integration in the smart home environment may create the following obstacles: first, blockchain mining is computationally intensive in nature, and this requires decent computing capabilities in the participating

devices; however, IoT devices have a heterogeneous computational power that may not be sufficient to mine blocks in a desirable amount of time. The second problem is that of data storage since blockchain is nothing but a distributed ledger stored locally to verify transactions; however, smart devices, such as various sensors, have a limited storage capacity. Finally, blockchain protocols also consume significant network capacity for internal communication, which may be undesirable for bandwidth-seeking smart devices.

Transportation Sector—In vehicle-to-vehicle (V2V) systems, vehicles communicate information with other vehicles. In vehicle-to-infrastructure (V2I) systems, vehicles communicate with the road network infrastructure to improve the safety and efficiency of the vehicle transportation infrastructure. Early examples of such systems are web-enabled tools such as Waze, a tool that provides real-time traffic conditions based on users, speed, and vehicle location. Even simple systems, such as automatic toll collection, stem from the idea of V2I. Futuristic examples are systems in which a car communicates its position on the road with other cars around it to prevent collisions, or for that vehicle to communicate with the transportation infrastructure, such as traffic signals, to provide better information on arriving vehicles to help it better manage traffic. In a global sense, each vehicle, traffic controller, and piece of road infrastructure can potentially become an IoT device, and each piece is connected to and can communicate through the Internet. The integrity of these devices must be ensured; the data that they produce will be critical. Blockchain technology can be applied to this area by confirming that a vehicle's ID is what it says it is, which is done by tracking its location on the road network. This would also prevent spoofed vehicles from tampering with or maliciously affecting automated systems. In 2014, Israeli students spoofed the Waze transportation application to report heavy congestion on a road when there was no traffic [6]. The students did this by reporting data on fake vehicles on the network at a particular road segment. A blockchain solution could have helped to autonomously manage vehicles IDs and their movements over the road network, and this would not allow systems such as Waze to account for vehicles that have just dropped onto particular segments. Moreover, the transportation infrastructure can ensure the integrity of vehicle data to help the autonomous systems make decisions based on a level of confidence in the data, backed by the knowledge of the valid history of the data.

Energy Sector—Blockchain technology has been proposed for use in the electric sector in the following application areas:

- Transactive energy—support distributed energy resource (DER) and its interaction with DER management systems (DERMS)
- eMobility—ability to transact energy charging at stations in multiple service territories
- Customer contracts—removing the middleman from the retail energy market

Blockchain technology can be used to provide supply chain security for the electric sector. Utilities are constantly installing new operations technology (OT) equipment and updating existing software and firmware in control system devices. One problem with this method is ensuring the integrity of the software and/or firmware. Some vendors use a digital signature when they distribute software and/or firmware updates, but this does not address the initial deployment. In addition, some vendors use a hash (typically MD5) as an integrity check. MD5 is not technically secure because it has a collision problem. That is, the same hash value can be computed on two different messages. Blockchain technology can be used to ensure software and/or firmware integrity in the electric sector security supply chain.

Consumer Electronics—Blockchain technology will impact the Consumer Electronics (CE) industry by providing cyber supply-chain provenance [7], where the customers as well as providers expect transparency for product information and delivery [8, 9]. Blockchain technology can mitigate cyber supply-chain risks for the CE industry by providing open access to the processes of planning, implementing, and controlling the movement of materials and finished goods to end users. Developing techniques and tools to provide provenance assurance are the top priority for addressing cyber supply-chain risks in the CE industry such as counterfeits, unauthorized production, tampering, theft, insertion of malicious software and hardware, as well as poor manufacturing and development practices. The globalization of the cyber supply chain has resulted in software and firmware being developed by offshore enterprises and has resulted in tremendous savings for the electronic data systems (EDS) sector. However, the dependency on third-party services has resulted in more maliciousness across the stages of the cyber supply chain. Specifically, there is a need for tools or technologies that can adequately address the risks involved in supply-chain processes, sourcing, third-party vendor management (every actor that has physical or virtual access to software code and/or systems), acquisition of compromised software or hardware purchases from suppliers, embedded malware in hardware or counterfeit hardware, and third-party data storage or data aggregators. Solutions exist, such as side-channel fingerprinting, reverse engineering, and formal methods, which are mostly deployed at the chip level to detect the presence of counterfeit chips. However, these methods cannot be scaled to protect the whole cyber supply chain. Thus, there is a need for blockchain-based methodology to maintain provenance across the supply-chain stages, as depicted in Figure 1.3. Radical evolution of IoT technology also attracts the majority of the CE industry to operate over cloud infrastructure [10]; thus, building a data provenance system will preserve transaction integrity and prevent malicious activities by alerting the users in real time.

Medical Sector—The recent influx of wearable medical devices promises to bring rich dividends to healthcare stakeholders. Wearable medical devices are networked computing devices equipped with sensors to track the patient's vital signs and physical activities. The data and the analytics can also be linked

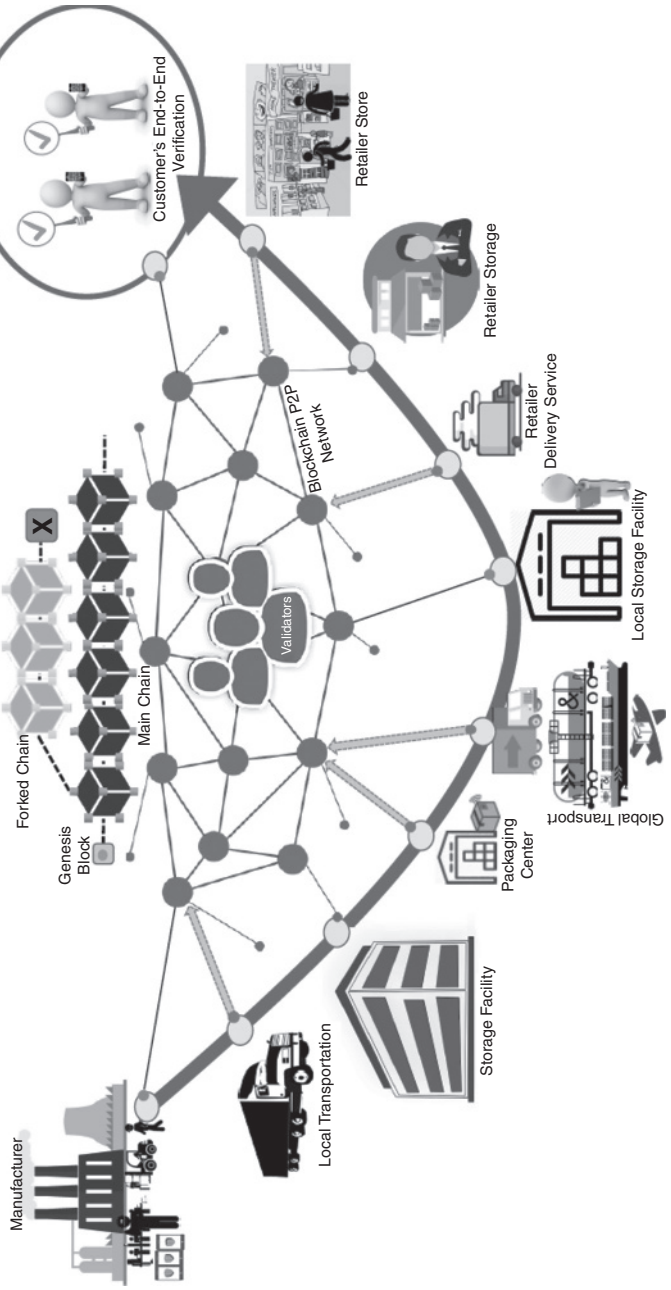

Figure 1.3 Supply-chain provenance overview.

to Electronic Health Records (EHR), which can benefit patients by helping them to monitor their personal health, aid doctors in prescribing personalized medicine, and allow insurance providers to gain insight into the cost of providing medical care.

However, due to security and privacy concerns, it has been reported that medical device manufacturers have only instrumented 20–30% of their networked devices to be used in hospitals. There have been several vulnerabilities reported with medical devices. For instance, ICS-CERT reported that Hospira's Symbiq drug infusion pump [11], used by medical facilities to automatically administer doses of medication to patients based on the amount specified by the caretaker, is vulnerable. The vulnerability allows an attacker to change doses of the prescribed medicine and impact patient safety. In 2017, the US Food and Drug Administration (FDA) reported vulnerabilities in St. Jude Medical's heart devices [12]. It is obvious that connected medical devices are here to stay, and the likelihood for compromising medical devices will grow exponentially. The current cyber-security solutions for identity management are inefficient and lack the ability to immediately track failure and accountability.

In addition to compromised medical devices, there are several privacy concerns with health data collected from both wearable devices and EHR systems. Patients are concerned about the lack of transparency in which a healthcare stakeholder has access to their data and how their data are used. Current healthcare cyber-security solutions focus on improving data providers' responsibilities to detect data disclosure activities; however, it is equally important to protect data access and provide immediate notifications of improper data disclosure risks. In addition, over 300 EHR systems use a centralized architecture that is prone to single point of failure and suffers from lack of interoperability that results in the absence of a holistic and thorough view of personal health. It is reported that 62% of insured adults rely on their doctors to manage their health records [13], which limits their ability to interact with healthcare providers other than their primary doctor. Moreover, even though health providers are supposed to follow rules or laws, such as HIPAA (Health Insurance Portability and Accountability Act of 1996), there are still many entities that are not covered by any laws. Therefore, it is crucial that any provider with access to data should be accountable for their operations on the data, and any operation on the data needs to be audited.

Blockchain's capability to capture data provenance will facilitate secure tracking of medical devices from production to ongoing use. The provenance information encoded in the blockchain provides immutable and reliable workflow with a trusted ground truth. The ground truth can be used for transparent traceability and accountability when any device malfunctions, accidentally or as a result of a security attack. The capability will also be useful for autonomous monitoring and preventive maintenance of medical devices. As compared to existing cyber-defense solutions, blockchain's distributed consensus protocols, cryptography techniques, and decentralized control will reduce cyber threat

risks for medical devices. Other benefits include streamlining the secure tracking of medical devices, cost savings, and improving patient privacy by secure and targeted access to patient data.

Blockchain relies on pseudoanonymity (replacing names with identifiers) and public key infrastructure (PKI) to maintain the privacy of users. The workshop [14] co-conducted by the Office of the National Coordinator (ONC) for Health IT and the National Institute for Standards and Technology (NIST) focused on blockchain usage in healthcare and research, aiming to clarify the implications of blockchain as an infrastructure for healthcare use cases including privacy preservation for predictive modeling, increasing interoperability between institutions on a large scale, immutability of health records, health insurance claim process improvement, health information exchange, healthcare delivery models with artificial intelligence, identity management, monetization strategies, and data provenance requirements.

Data Provenance in the Cloud—Assured data provenance in cloud computing is needed to keep track of data transactions generated by data operations in the cloud and detect malicious activities. The current state-of-the-art data provenance technologies involve comparing logged data generated by execution of software on physical or virtual resources. However, they fail to detect integrity violations and are typically conducted in a private setting to allow better ownership of assets. Also, the process is not scalable to federated cloud environments, is costly, and lacks transparency. Hence, there is a strong need to develop a data provenance framework for the cloud computing environment, where multiple representatives or virtual stakeholders can participate in maintaining transparent and immutable provenance information. Blockchain technology, where data are stored in a public, distributed and immutable ledger and maintained by a decentralized network of computing nodes, provides a decentralized and permanent record-keeping capability, which is critical for data provenance and access control in cloud data protection.

1.1.3 Blockchain Military Cyber Operations Use Cases

Blockchain as a distributed ledger system provides many features and functionalities that are needed for cyber operations, such as auditing of historical information, assured data provenance, guaranteed variability of integrity violations of historical data, and auditing of contents from tampering. Besides, blockchain has both cost effectiveness merits as well as transparency features, making it an appealing system for military cyber operations.

- Generation of cyber assets—Blockchain can be utilized to generate cyber assets that will enable applications that rely on direct interaction between customers and assets. The blockchain system can aid in assuring the processes of issuance, transaction processing, and housing of cyber assets and identities.

- Transfer of ownership of cyber assets—The blockchain system allows transfer of cyber assets between owners by leveraging the property of blockchain so that once a transaction is guaranteed, it cannot be reversed. Any changes will have to be appended and will not result in change of an already validated transaction, thereby ensuring nonreversibility of transfer of ownership.
- Transparent and assured data provenance—Every operation on the cyber asset is encoded in the blockchain transaction using a publicly available and immutable ledger. The blockchain system ensures that provenance of every operation on the cyber asset is recorded and traceable.
- Verifiability and audit—The distributed ledger keeps track of transactions pertaining to creation and transfer of cyber assets. The tamper-resistant property of the ledger facilitates variability and audit of operations.

Military Cyber Operations—Ensuring traceable and tamper-evident accountability and auditability of command and control, logistics, and other critical mission data among international partners is paramount as our future operations involve the convergence of multiple domains and a heavily contested cyberspace. Centralized or homogenous information systems and databases must evolve distributed, disintermediated, and secure capabilities. As such, trust with respect to operations involving international entities must not be rooted in one single entity. Trust must be decentralized and built around robust, innovative cryptographic paradigms transcending the traditional PKI typically utilized in most homogenous enterprises.

An innovative, distributed trust and identity management mechanism is a crucial enabling technology assuring identification, authentication, and authorization in a way that would further allow disintermediated accountability and auditability. Emerging blockchain and distributed ledger technologies as a whole demonstrate the potential of a truly distributed and disintermediated mechanism for accountability and auditability. The current production application of cryptocurrencies has already created unprecedented accountability and auditability in a way that disrupts traditional fiat currencies and disintermediates the way people are able to transact internationally via pseudonymous identity management via wallets in permissionless and public blockchain implementations.

The nuances of disintermediated international partnerships and information exchange involve some mutually exclusive research and development challenges distinct from the permissionless and public implementations of blockchain. For the level of identification, authentication, authorization, accountability, and auditability challenges that encompasses overall integrity concerns for international operations, even more robust distributed identity management mechanisms must be investigated. Furthermore, the underlying practice of consensus must vary from the latency-tolerant and performance-heavy implementations observed in cryptocurrencies. Most importantly, the distributed ledger technology mechanism may be private and permissioned;

however, it must also not be vested in one single vendor's implementation in a way that inadvertently and ironically intermediates and centralizes information exchanges. We must thoroughly assess and demand interoperability and standardization among blockchain implementations for mission data, which further enables the disintermediated accountability and auditability required in our increasingly complex international operations in multiple domains.

1.1.4 Blockchain Challenges

Blockchain does have the potential to address cyber security issues in distributed systems. However, blockchain cannot be considered a panacea to address all cyber security concerns. Blockchain does have inherent capabilities to address integrity violations. However, assurance of confidentiality, availability, and authenticity is not guaranteed by blockchain and will require integration with several security solutions. Organizations that are considering blockchain to address cyber security issues should carefully assess whether the technology is a good fit. Specifically, the below outline is a good start to identify whether the organization does need blockchain to replace the existing solution [15]:

- Do you have concerns regarding the ability of the centralized database to withstand failures?
- Are there multiple stakeholders responsible for modifying the contents of the database?
- Do the multiple parties operate under differing trust domains?
- Are there clear defined rules to control data input?
- Is there a value proposition in having validators in the form of consensus?

Once the need for distributed blockchain has been justified, the next steps are to identify the blockchain solution that's a good fit for the organization. There are several aspects that need to be addressed, such as the type of data encoded in the blockchain transactions, frequency of transactions, the infrastructure used to store the blockchain (public or permissioned), key management system, number of validators, bootstrap time, ability of smart contracts to learn rules dynamically, attack surfaces in the blockchain solution, etc. For instance, for the healthcare sector, the type of data stored in the blockchain needs to be carefully identified as any sensitive information can be subjected to confidentiality attacks. Any organization that would like to ensure that activities of participants in the blockchain are compliant would prefer permissioned blockchain [16].

Below is a summary of key research challenges that need to be addressed for realizing a practical blockchain solution.

Scalability—Bitcoin's current implementations are not scalable due to the fact that it takes 10 minutes or longer to confirm transactions and seven transactions is the maximum throughput that can be achieved. There is a need for

fundamental research to develop a scalable blockchain platform. Prior to developing a scalable blockchain platform, it is imperative to not only define scalability in the context of blockchain, but also identify metrics to quantify scalability. There have been efforts to improve the scalability in blockchain by modifying parameters such as block size and block intervals. However, these efforts to achieve scaling through reparameterization alone can only realize limited benefits and do not address network performance issues. Network performance is exacerbated due to blockchain's p2p overlay network protocol, degree of decentralization, and number of peers in the network. The throughput of blockchain depends on the throughput of the overlay network that determines the rate at which blocks propagate and the percentage of nodes involved in the exchange of blocks in a given time interval. For example, if the transaction rate reaches 80% of the throughput, it is quite possible that 10–20% of the p2p nodes will not be able to render services and reduce effective network mining power.

There is a need to develop new architectures for blockchain to ensure sufficient scalability without sacrificing decentralization. The architecture should involve protocol design strategies across several layers, namely network, consensus, and storage. There is also a need to identify and measure scalability metrics such as throughput, latency, bootstrap time, storage, cost of confirmed transaction, fairness, and network utilization. The architecture will also need to be designed to address issues such as, "Does exploitation of system parameters to improve scalability sacrifice security properties?" and "What is the degree of resilience of the system during a cyberattack?"

Network Layer—The objective of the network layer in the blockchain architecture is to provide an effective mechanism to propagate transaction messages. The network layer ensures that messages from any participant can be transmitted to all the nodes in the blockchain network. However, the network layer does not operate in full broadcast mode and nodes exchange messages that contain validated transactions. And, in most current implementations of the blockchain network, the network is heavily underutilized and limits throughput. Thus, the network layer in blockchain is a bottleneck in the processing of transactions.

Consensus Layer—The consensus layer is responsible for validating transactions and uses the network layer to deliver messages and record the transactions in the distributed ledger. The consensus protocols include proof of work (PoW), proof of stake (PoS), and byzantine fault tolerance. Traditional blockchain technology relies heavily on the underlying PoW mechanism to achieve consensus in the decentralized system where the miner has to spend its computational power to solve the cryptopuzzle so as to successfully include its block in the blockchain. With such an approach, miners opt for various specialized hardware to achieve their computational ability. The eventual goal of the miners is to win the block-adding race so that they can be rewarded, and a significant amount of energy is required to do so. For a simple example, if we consider the case of Bitcoin's blockchain, the miners compete to get the reward of 25 Bitcoin, which is worth approximately $20,000 and is freshly minted for the

winning miner every 10 minutes. Thus, the amount of reward per second is $33.30, and if we assume the rate of industrial electricity is $0.01/KWhr, then we can approximately state that Bitcoin miners use energy of 1100 MW per second. This substantial quantity is spent to reach consensus using the PoW approach and most of it is used in computing the irreversible SHA256 hashing function. Since the value of direct incentives will diminish eventually, the critical question of "how will the PoW miners be motivated to mine?" has to be addressed so as to smoothly run the consensus process. PoS consensus protocol is interestingly attractive; it provides block inclusion decision-making power to those entities that have stakes in the system irrespective of the blockchain's length or history of the public ledger. The principal motivation behind this scheme is to place the power of leader election in the blockchain update process into the hands of the stakeholders. This is done to ensure that the security of the system will be maintained while the members' stakes are at risk. Roughly speaking, this approach is similar to the PoW consensus except the computational part. Hence, a stakeholder's chances to extend the blockchain by including its own block depend proportionately on the amount of stake it has in the system.

There is a need for developing a customized consensus engine that will not require participants to make significant investments in computation and will balance the tradeoff between the number of transactions processed, transaction validation time, incentives, and security rules set by participators. The customized consensus engine will choose the optimal combination of consensus protocols to achieve the aforementioned objectives.

Privacy—Permissioned blockchain platforms, such as Hyperledger Fabric and JP Morgan's Quorum, claim that privacy is a goal; but the way they achieve it is actually quite limited. These systems consist of several validating nodes, each of which sees the entire transaction log in plaintext. That is, while the systems are designed to provide availability/consistency even when some of the nodes fail, they cannot guarantee privacy if one of the nodes suffers a data breach. Some permissioned blockchain platforms offer a feature where you can create a "private channel" comprising just a subset of these nodes; however, among this subset, it is still the case that any data breach would leak the transaction data and then the private channel cannot interact with the other channels. With the existing systems, there is an inherent tradeoff between resilience and expressiveness on the one hand, and privacy on the other.

There are a variety of technical approaches that can provide better operating points. These include threshold cryptography/multi-party computation, zero knowledge proofs, and homomorphic encryption. The theory for these approaches is well established in general, but in concrete terms it is a great open challenge to (a) find efficient algorithms for applications of interest, and (b) integrate with existing systems. So, focusing on building better privacy mechanisms could be a well-motivated and technically interesting challenge to add.

Security—Despite the advantages of using blockchain for distributed systems security, there have been numerous instances of reported security risks

associated with this technology [17–19]. In 2016, it was reported that an adversary was able to withdraw $50 million from "The DAO", a decentralized autonomous organization that operates on blockchain-based smart contracts [20]. In June 2017, Bitfinex reported a distributed denial-of-service (DDoS) attack that led to a temporary suspension. Several exchanges of Bitcoin and Ethereum (a blockchain-based distributed computing platform) have also suffered from DDoS attacks and Domain Name System (DNS) attacks frequently, hampering service availability to the users.

The attack surfaces in blockchain can be broadly classified into the following three main categories: (i) Threats associated with the techniques employed for creating/maintaining the distributed ledger (e.g. blockchain forks, stale blocks, orphaned blocks, etc.), (ii) threats to the blockchain system's underlying network infrastructures (e.g. attacks on consensus protocols that cause delays, decreased throughput, inconsistencies, DDoS attack, DNS attacks, Fork After Withholding [FAW] attacks, etc.), and (iii) threats associated with frontend/backend applications integrated with blockchain technology (stealing of private keys, attacks on certificate authorities, attacks on membership services in permissioned blockchain, blockchain ingestion, double spending, wallet theft, etc.).

1.2 Overview of the Book

The book has encompassing themes that drive the individual contributions including the broad theme of blockchain-based secure data management and storage for cloud and IoT; data provenance in cloud storage; secure IoT model, auditing architecture, and application of blockchain in military and healthcare domain; and empirical validation of permissioned blockchain platforms. It will synthesize a mix of earlier works (on topics including data provenance in blockcloud and blockcloud security analysis) as well as newer, cutting-edge research findings that promise to attract strong interest (on topics including invariant-based supply chain protection, information-sharing framework, and trustworthy information federation). The contributions address security and privacy concerns in blockchain in key areas, including the following: preventing digital currency miners from launching attacks against mining pools, empirical analysis of the attack surface of blockchain, countering double spending in blockchains, security analysis of blockchain consensus protocols, and privacy in permissioned blockchain platforms.

1.2.1 Chapter 2: Distributed Consensus Protocols and Algorithms

Fault-tolerant consensus has been extensively studied in the context of distributed computing. By regulating the dissemination of information within the

network of distributed processors, a fault-tolerant consensus mechanism guarantees that all processors agree on common data values and perform the same course of action in response to a service request, in spite of faulty processors and unreliable communication links. This consensus guarantee is crucial to normal functioning of the distributed computing system. Similarly, as a realization of a distributed system, blockchain requires a consensus protocol that assures all nodes in the p2p network agree on a single chain of transaction history (or "public ledger") given the adverse effect of malfunctioning nodes and unpredictable network conditions. At the time of writing, there are more than a hundred blockchain primitives in the cryptocurrency market, embodying more than 10 classes of consensus protocols. Aiming at the fundamentals of distributed consensus, this chapter provides an overview of topics ranging from classic fault-tolerant consensus in distributed computing to the current stage of blockchain consensus protocols. Essentially, analysis of consensus performance will be aided by mathematical modeling. In this chapter, we first provide the basics of fault-tolerant consensus in distributed computing in a succinct manner. Next, we conduct formal analysis of the Nakamoto protocol—the pioneering PoW blockchain consensus protocol for Bitcoin. We will also present several emerging non-PoW blockchain consensus protocols and their application scenarios. We will provide a qualitative evaluation and comparison over the aforementioned blockchain consensus protocols.

1.2.2 Chapter 3: Overview of Attack Surfaces in Blockchain

In this chapter, we explore the attack surface of blockchains and the possible ways in which this technology can be compromised. Toward this goal, we attribute attack viability in the attack surface to (i) blockchain cryptographic constructs, (ii) the distributed architecture of the systems using blockchain, and (iii) the blockchain application context. For each of those contributing factors, we outline several attacks, including selfish mining and associated peer behaviors, 51% attack, DNS attacks, DDoS attacks, equivocation, consensus delay (due to selfish behavior or DDoS attacks), blockchain forking, orphaned and stale blocks, block ingestion, wallet thefts, and privacy attacks. We then explore the causal relationship between these attacks and show how one fraudulent activity can lead to the possibility of other attacks. A secondary contribution of this work is outlining effective defense measures taken by blockchain technology or proposed by researchers to mitigate the effects of these attacks and patching the vulnerabilities in blockchains.

1.2.3 Chapter 4: Data Provenance in Cloud Storage with Blockchain

This chapter broadly covers the topic of data provenance in cloud storage and key challenges. It presents ProvChain, a blockchain platform that provides

assured data provenance, which achieves the following objectives: (i) real-time cloud data provenance, (ii) tamper-proof environment, (iii) enhanced privacy preservation, and (iv) provenance data validation. A detailed design to achieve the aforementioned objectives within ProvChain is given. An implementation of ProvChain on ownCloud, an open source cloud storage service, is provided. Then, it presents detailed evaluation results to demonstrate the effectiveness of ProvChain to provide assured data provenance with the desired privacy and availability in cloud storage platforms. Finally, it describes and offers some solutions to the challenge of processing graph data in the cloud.

1.2.4 Chapter 5: Blockchain-based Solution to Automotive Security and Privacy

Smart vehicles are increasingly being connected to other vehicles in close proximity and to roadside infrastructures, for example, traffic lights and overhead displays at motorways, and more generally to the internet, thus making the vehicles part of the IoT. This high degree of connectivity introduces new, sophisticated personalized services for smart vehicle owners as well as for vehicle manufacturers, suppliers, and service providers (SPs) such as insurance companies. However, this high degree of connectivity makes smart vehicles highly vulnerable to security threats and raises various privacy concerns.

In this chapter, we present a decentralized privacy-preserving and secure blockchain-based architecture for a smart vehicle ecosystem. Smart vehicles, Original Equipment Manufacturer (OEM) (i.e. car manufacturers), and other SPs jointly form an overlay network where they can communicate with each other. Nodes in the overlay are clustered and only the cluster heads are responsible for managing the blockchain and performing its core functions. These nodes are known as overlay block managers (OBMs). Transactions are broadcast to and verified by the OBMs, thus eliminating the need for a central broker. To protect user privacy, each vehicle is equipped with in-vehicle storage to store privacy-sensitive data, for example, location traces. The vehicle owner defines which data (and the granularity) is traded to third parties for beneficial services and which data should be stored in the in-vehicle storage. Consequently, the owner has greater control over the exchanged data.

All transactions (i.e. communications) in the network are encrypted using asymmetric encryption. Nodes are authenticated using their Private Keys (PKs). Strong communication security and authentication introduced by blockchain mitigates the risk of the vehicle being remotely hacked and thus increases the safety of the passengers. We undertake a qualitative analysis on the security and privacy of the proposed architecture. Multiple possible attacks and the protection methods employed by our framework are discussed to show the resilience of our framework against them. We develop an implementation as a proof of concept to demonstrate the applicability of our approach and to analyze its packet and delay overheads.

1.2.5 Chapter 6: Blockchain-based Dynamic Key Management for IoT-Transportation Security Protection

IoT is the next generation platform that aims to maximize the connection between the cyber platform and the physical world including, but not limited to, vehicles, infrastructures, home sensors, smart medical systems, and wearable electronics. However, security is still the main concern in IoT environments. Even though significant developments in security assurance schemes have taken place over the past few years in the area of communication engineering, application layer security, especially cross-domain and cross-scenario (heterogeneity) security schemes, are still an open topic for research. Moreover, blockchain technology provides a feasible solution for these security challenges. This chapter describes a blockchain-based IoT platform for transportation security, specifically, a blockchain-based solution that suits one of the emerging IoT use cases—Vehicular Communication Systems (VCS), one of the most important IoT components and a subsystem of Intelligent Transportation Systems (ITS). The solution mainly adapts to simplify the distributed key management in heterogeneous VCS domains. A dynamic transaction collection period is proposed to minimize the key transfer time during vehicle handover. Furthermore, potential developments in the privacy prevention area are demonstrated.

1.2.6 Chapter 7: Blockchain-enabled Information Sharing Framework for Cybersecurity

This chapter examines the design, development, and evaluation process established for a blockchain-based information sharing (BIS) framework. The BIS at its core offers a mechanism for confidential information and infrastructure protection from future cyberattacks leveraging on blockchain technology. Blockchain is the concept used in the Bitcoin system, currently explored to provide transparent p2p transactions in multiple domains. In BIS, multiple organizations share security-related information, while preserving their privacy, in a bid to jointly protect their cyberspace. The bigger picture here is to collect high-resolution cyberattack information from multiple organizations with the organizations having no specific knowledge about the usage of other organizations' data or exposure of information about a particular company's cybersecurity attacks. BIS offers a decentralized approach for blockchain with transactions being digitally signed to ensure identification of legitimate organizations. This allows the detection of adversaries posing as legitimate users to learn from the public ledger with hashed pointers. Furthermore, the activities of nonparticipating users of BIS for security attacks and defenses are analyzed using a Stackelberg game.

Also included in the chapter is blockchain-based protocol for cyber threat information sharing, an overview of integrated information-sharing

framework (iShare), blockchain protocol that protects confidential information from adversarial threats, and a game-theoretic approach to analyze security attacks on iShare.

1.2.7 Chapter 8: Blockcloud Security Analysis

In this chapter, we focus on security analysis of blockchain-based data provenance in cloud. Blockchain's public and distributed p2p ledger capability benefits cloud computing services, which require functions such as assured data provenance, auditing, management of digital assets, and distributed consensus. Blockchain's underlying consensus mechanism allows the building of a tamper-proof environment where transactions on any digital assets are verified by a set of authentic participants or miners. However, achieving consensus demands computational power from the miners in exchange for a handsome reward. Therefore, greedy miners always try to exploit the system by augmenting their mining power. In this chapter, we first discuss blockchain's capability in providing assured data provenance in cloud and present vulnerabilities in blockchain cloud. We model the block withholding (BWH) attack in a blockchain cloud considering distinct pool reward mechanisms. BWH attack provides a rogue miner ample resources in the blockchain cloud for disrupting honest miners' mining efforts, which was verified through simulations.

1.2.8 Chapter 9: Security and Privacy of Permissioned and Permissionless Blockchain

Most analysis of blockchain protocols we have seen so far, ranging from traditional consensus protocols like Paxos and PBFT to Nakamoto consensus, all rely on the "majority honest" assumption, where we assume a majority of the parties follow the protocol correctly. But why should we be willing to assume that any of the peers will be honest and run the protocol exactly? In this chapter, we discuss two security design considerations and apply them to both permissioned and permissionless models. The first design approach is committee selection, by which a large population of participants are winnowed down into a small, fairly sampled subset, where the attacker does not have much presence in the committee. This design approach applies equally well in both permissionless and permissioned blockchains since it can improve performance versus having the entire population active. The second issue we discuss is privacy. Blockchain applications often need to provide privacy guarantees for users, for example, if they involve sensitive information about financial transactions or about the real-time location of Internet of Things devices. Here, cryptography can be employed. If we have a high degree of trust in the peers, as in a permissioned setting, secret sharing is a natural approach since we assume a majority of the peers will not be breached. On the other hand, in a context with

less trust, zero-knowledge proofs allow clients to prevent any of the peers from seeing protected data.

1.2.9 Chapter 10: Shocking Public Blockchains' Memory with Unconfirmed Transactions—New DDoS Attacks and Countermeasures

In 2017, blockchain-based systems have seen a rise in their value. Consequently, they have attracted several forms of denial-of-service attacks, and their attack surface is being widely explored in the fields of security and distributed systems. In this chapter, we present a new form of attack that can be carried out on the memory pools (mempools) of blockchain systems in general. Towards that end, we study such an attack and explore its effects on transaction fee structures of legitimate users. We also propose countermeasures to contain such an attack. Our countermeasures include fee-based and age-based designs, which optimize the mempool size and help in countering the effects of this attack. We further evaluate our designs by simulations and analyze their usefulness in varying attack conditions. Our analyses can be extended to a wide variety of blockchain systems using proof concepts, where fees are provided as an incentive for participation.

1.2.10 Chapter 11: Preventing Digital Currency Miners From Launching Attacks Against Mining Pools by a Reputation-Based Paradigm

The mining process in blockchain is very resource intensive; therefore, miners form coalitions to verify each block of transactions in return for a reward where only the first coalition that accomplishes the PoW will be rewarded. This leads to intense competitions among miners and, consequently, dishonest mining strategies such as BWH attack, selfish mining, eclipse attack, and stubborn mining, to name a few. As a result, it is necessary to regulate the mining process to make miners accountable for any dishonest mining behavior. In this chapter, we propose a new reputation-based framework for PoW computation in blockchain in which miners are not only incentivized to conduct honest mining, but also disincentivized for committing any malicious activities against other mining pools. We first illustrate the architecture of our reputation-based paradigm, explain how the miners are rewarded or penalized in our model, and subsequently, we provide game-theoretical analyses to show how this new framework encourages the miners to avoid dishonest mining strategies. In our setting, a mining game is repeatedly played among a set of pool managers and miners where the reputation of each miner or mining ally is continuously measured. At each round of the game, the pool managers send invitations only to a subset of miners based on a nonuniform probability distribution defined by the miners' reputation values. We show that by using our proposed solution concept, honest mining attains Nash equilibrium in our setting. In other words, it will not be in the best interest of the miners to employ dishonest mining

strategies even by gaining a short-term utility. This is due to the consideration of a long-term utility in our model and its impact on the miners' utilities over time.

1.2.11 Chapter 12: Private Blockchain Configurations for Improved IoT Security

Blockchain has the potential to provide secure device authentication, trusted event logs, and interoperability for IoT. There are several blockchain platforms to realize the aforementioned requirements; however, there is no consensus on the optimal blockchain platform and the necessary configuration that is applicable to IoT deployments. In this chapter, we focus on the methodology to configure the blockchain technology to meet the security requirements of IoT.

We describe, implement, and compare two possible private blockchain configuration strategies—blockchain-enabled gateways and blockchain-enabled end devices. Test use cases for both strategies are implemented on a network of Raspberry Pi devices using the popular Ethereum and Hyperledger Fabric blockchain frameworks, respectively. We show that while the more popular blockchain gateway approach is better suited to the current architecture and computational requirements of leading blockchain frameworks, the blockchain-enabled end device approach is technically feasible and highly promising, enabling more trustworthy data collection and complex device management strategies. The chapter focuses on the different roles that IoT devices can play within a blockchain deployment as well as the security guarantees that the blockchain can provide an IoT device in a certain role. We also identify system functionality and cyber security guarantees in scenarios wherein IoT devices are configured as full nodes of the blockchain.

1.2.12 Chapter 13: Blockchain Evaluation Platform

In the final chapter, we provide a systematic methodology to simulate, test, and evaluate blockchain platforms. The methodology described in this chapter will provide empirical means to evaluate the efficacy of the approaches presented in preceding chapters. This chapter strives to provide some insight into how blockchain theoretical models can be simulated and tested on a practical platform. The chapter will describe the development of a blockchain simulator that can be used to conduct performance and security evaluation. We will also provide insights into blockchain implementation within a Hyperledger Fabric, an open-sourced blockchain application and toolset managed by the Linux Foundation. The Hyperledger Fabric example is a fully capable blockchain platform that can be modified to work in a real environment. Both examples can be scaled to test the performance of protocols and theories as needed, which makes them both good examples to use. The software code will be available to

the reader via an online repository and can be downloaded and run generally on any system with some open-sourced software requirements.

References

1 J. Kelly and A. Williams, "Forty big banks test blockchain-based bond trading system." [Online]. Reuters, March 3, 2016. Available: https://www.reuters.com/article/banking-blockchain-bonds/forty-big-banks-test-blockchain-based-bond-trading-system-idUSL8N16A30H.
2 I. Kar, "Estonian citizens will soon have the world's most hack-proof health-care records." [Online]. Quartz, March 3, 2016. Available:https://qz.com/628889/this-eastern-european-country-is-moving-its-health-records-to-the-blockchain/.
3 S. Lacey, "The energy blockchain: How bitcoin could be a catalyst for the distributed grid." [Online]. Green Tech Media, February 26, 2016. Available: https://www.greentechmedia.com/articles/read/the-energy-blockchain-could-bitcoin-be-a-catalyst-for-the-distributed-grid#gs.4cTkNvw.
4 D. Oparah, "3 ways that the blockchain will change the real estate market." [Online]. TechCrunch, February 7, 2016. Available: https://techcrunch.com/2016/02/06/3-ways-that-blockchain-will-change-the-real-estate-market/.
5 A. Mizrahi, "A blockchain-based property ownership recording system." [Online]. 2015. Available: https://chromaway.com/papers/A-blockchain-based-property-registry.pdf.
6 N. Tufnell, "Students hack Waze, send in army of traffic bots." [Online]. Wired, March 25, 2014. Available: http://www.wired.co.uk/article/waze-hacked-fake-traffic-jam.
7 K. Toyoda, P. T. Mathiopoulos, I. Sasase, and T. Ohtsuki, "A novel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain," *IEEE Access*, vol. 5, no. 99, pp. 17465–17477, 2017.
8 J. H. Lee and M. Pilkington, "How the blockchain revolution will reshape the consumer electronics industry [future directions]," *IEEE Consumer Electronics Magazine*, vol. 6, no. 3, pp. 19–23, July 2017.
9 D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and C. Yang, "The blockchain as a decentralized security framework," *IEEE Consumer Electronics Magazine*, vol. 7, no. 2, pp. 18–21, 2018.
10 J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (IoT): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.
11 M. Mellen, "Tip of the Iceberg: FDA's alert to unplug Hospira's drug infusion pumps from clinical networks." [Online]. August 4, 2015. Available: https://researchcenter.paloaltonetworks.com/2015/08/tip-of-the-iceberg-fdas-alert-to-unplug-hospiras-drug-infusion-pumps-from-clinical-networks/.

12 A. Stark, "FDA issues safety communication on availability of firmware update to address cybersecurity vulnerabilities identified in Abbott's (formerly St. Jude Medical's) implantable cardiac pacemakers." [Online]. 2018. Available: https://www.fda.gov/NewsEvents/Newsroom/FDAInBrief/ucm573853.htm.

13 2016 connected patient report, https://www.salesforce.com/assets/pdf/industries/2016-state-of-the-connected-patient-pr.pdf.

14 C. Ryan, "Use of blockchain in healthcare and research workshop." [Online]. The Office of the National Coordinator (ONC) for Health IT, the National Institute for Standards and Technology (NIST), November 2016. Available: https://oncprojectracking.healthit.gov/wiki/display/TechLabI/Use+of+Blockchain+in+Healthcare+and+Research+Workshop.

15 G. Greenspan, "Avoiding the pointless blockchain project." [Online]. November 22, 2015. Available: http://www.multichain.com/blog/2015/11/avoiding-pointless-blockchain-project/.

16 S. Friedman, "Before the blockchain: 4 questions to answer." [Online]. June 25, 2018. Available: https://gcn.com/articles/2018/06/25/blockchain-questions.aspx.

17 X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems." [Online]. 2018. Available: http://arxiv.org/abs/1802.06993.

18 I.-C. Lin and T.-C. Liao, "A survey of blockchain security issues and challenges." *IJ Network Security*, vol. 19, no. 5, pp. 653–659, 2017.

19 N. Atzei, M. Bartoletti, and T. Cimoli, "A survey of attacks on Ethereum smart contracts (SoK)," in Proceedings of the 6th International Conference on Principles of Security and Trust—volume 10204, Sweden, 2017, pp. 164–186.

20 D. Siegel, "Understanding the DAO attack." [Online]. June 25, 2016. Available: https://www.coindesk.com/understanding-dao-hack-journalists/.