Blockchain versus Database: A Critical Analysis

Mohammad Jabel Morshed Chowdhury*, Alan Colman*, Muhammad Ashad Kabir†, Jun Han* and Paul Sarda*

*School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, Australia
Email: {mjhchowdhury, acolman, jhan, 4949501}@swin.edu.au
†School of Computing and Mathematics Charles Sturt University, NSW, Australia
Email: akabir@csu.edu.au

Abstract—In recent times, Bitcoin has gained special attention both from industry and academia. The underlying technology that enables Bitcoin (more generally crypto-currency) is called blockchain. At the core of the blockchain technology is a data structure that keeps record of the transactions in the network. The special feature that distinguishes it from existing technology is its immutability of the stored records. To achieve immutability, it uses consensus and cryptographic mechanisms. As the data is stored in distributed nodes this technology is also termed as “Distributed Ledger Technology (DLT)”. As many researchers and practitioners are joining the hype of blockchain, some of them are raising the question about the fundamental difference between blockchain and traditional database and its real value or potential. In this paper, we present a critical analysis of both technologies based on a survey of the research literature where blockchain solutions are applied to various scenarios. Based on this analysis, we further develop a decision tree diagram that will help both practitioners and researchers to choose the appropriate technology for their use cases. Using our proposed decision tree we evaluate a sample of the existing works to see to what extent the blockchain solutions have been used appropriately in the relevant problem domains.

Index Terms—blockchain, database, distributed ledger technology.

I. INTRODUCTION

Blockchain technology aims at creating a decentralized environment where no third party is in control of the transactions and data [19]. In general, the blockchain is a time-stamped chain of blocks jointly maintained by all participating nodes. Blocks are basically containers that aggregate transactions. The blocks are chained together cryptographically: each block is digitally signed and ‘chained’ to the previous block by including that block’s hash value. New blocks can only be appended to the end of the chain, thus the blockchain provides an immutable data storage (existing transactions cannot be updated or deleted). For this reason, many systems built on the blockchain technology achieve secured distribution of digital assets among untrusted clients.

Blockchain has been used in several domains due to the benefits of distributed data storage and immutable audit trails. In healthcare, several approaches have been introduced in the field of Electronic Health Records (EHR) [2] [3] [4]. Due to the transparency of the technology, governments and businesses also try to apply the technology and harvest its benefits [5] [6] in applications such as electronic cash systems, business process management and supply chains [7]. In the emerging field of the Internet of Things (IoT), blockchain technology can be used in different scenarios and forms [8]. These include the management of privacy and security of IoT [9], as well as the development of new scenarios and business opportunities. The blockchain technology is also being applied in higher education [10]. Several higher education institutions have employed the blockchain technology to design different solutions and approaches for managing student records. In general, most of the blockchain solutions in education sector use the Bitcoin blockchain [11].

There is little doubt that the blockchain technology is still in the initial phase of Gartner’s Hype Cycle for emerging technologies [25]. Government and research organizations are pouring funding for blockchain related research. Recently, the European Commission has announced 30 billion on new investments in technology initiatives including those involving blockchain [12]. However, there is considerable debate in the community about the value of blockchain over a shared database [1]. For example, Narayanan contends that private blockchains are just another name for a shared database [13]. Others, like Greenspan, see several differences between private blockchains and SQL databases, from trust building to robustness [14].

While blockchain is a powerful technology but if it is indiscriminately applied to use cases without considering the strength and weakness of the technology, we will fail to realize the true potential of this technology. We have therefore conducted a scoping review [16] to understand how different researchers are using this technology. Our findings reveal that most of the exiting literatures focus on “how” the blockchain technology works and, to a lesser extent, on “what” (potential) applications and usages that business organizations can leverage. For its part, the “why” question, which focuses on the organizational motivations for adopting the blockchain technology, was scarcely discussed in the literature.

The main question we address with this paper is, whether or not the proposed applications of the blockchain technology as a solution to problems, have a well-founded basis. In answering the question above, we make a number of contributions,

1) We have conducted a scope review to see the trend in the volume and areas of research concerning blockchain in the last 5 years.
We have provide a comparative analysis between the blockchain and database technologies.

3) We propose a decision tree to check whether the use case in hand should use blockchain or not. If yes, what kind of blockchain technology should be used.

4) Using the decision tree, we have also done a preliminary assessment of the existing literature to gauge the extent to which the proposed uses of blockchain are justifiably in terms of the unique advantages that blockchain provides.

In section 2, we give an overview the properties of the blockchain technology and its various variants. Section 3 presents the scoping review of the blockchain technology. We discuss three popular use cases of blockchain in section 4. We present a comparison between blockchain and database in section 5. Section 6 presents and discusses a decision-making process (in the form of decision tree / flow chart) that helps to determine when to use which type of blockchain. Using this decision-making process, We assess the appropriateness of a sample of existing uses of blockchain in section 7. We conclude the paper in Section 8.

II. BLOCKCHAIN

Bitcoin [15], introduced in 2008, has emerged as the world’s first widely used digital currency and has been used in a wide range of applications. Interestingly, it is underpinned by a novel mechanism called the Distributed Ledger Technology (DLT), also known as the blockchain technology, providing its solid technical foundation.

Even though the terms blockchain and DLT are used interchangeably in the literature, there is a subtle difference between them which is worth highlighting. A blockchain is just an example of a particular type of ledger, and there are other types of ledger. When a ledger (including a blockchain) is distributed across a network, it can be regarded as a Distributed Ledger. For simplicity, we will also use the terms interchangeably.

In the last few years, blockchain has received widespread attention from the industry, the government and the academia alike and is regarded as one of the fundamental technologies to revolutionise the landscapes of several application domains.

At the centre of the blockchain technology is the blockchain itself. A blockchain is a ledger consisting of consecutive blocks chained together following a strict set of rules. The ledger is distributed and stored by the nodes of a P2P network where each block is created at a predefined interval in a decentralized fashion by means of a consensus algorithm. The consensus algorithm guarantees several data integrity related properties (discussed below) in blockchain. A blockchain exhibits several properties which make it a suitable candidate for several application domains. The properties are discussed below.

- **Distributed consensus on the chain state:** One of the crucial properties of any blockchain is its capability to achieve a distributed consensus on the state of the chain without being reliant on any trusted third party. This opens up the door of opportunities to build and utilize a system where every possible state or interaction is verifiable by the authorised entities.

- **Immutability and irreversibility of chain state:** Achieving a distributed consensus with the participation of a large number of nodes ensures that the chain state becomes practically immutable and irreversible after a certain period of time. This also applies to smart-contracts and hence enabling the deployment and execution of immutable computer programs.

- **Data (transaction) persistence:** Data in a blockchain is stored in a distributed fashion ensuring its persistency as long as there are participating nodes in the P2P network.

- **Data provenance:** The data storage process in any blockchain is facilitated by means of a mechanism called the transaction. Every transaction needs to be digitally signed using public key cryptography which ensures the authenticity of the source of data. Combining this with the immutability and irreversibility of a blockchain provides a strong non-repudiation instrument for any data in the blockchain.

- **Distributed data control:** A blockchain ensures that data in the chain are stored in a distributed manner that exhibits no single point of failure.

- **Accountability and transparency:** Since the state of the chain, along with every single interaction among participating entities, can be verified by any authorised entity, it promotes accountability and transparency.

III. SCOPING REVIEW

In this study we have taken from Arksey [16] O’Malleys [17] and Levac et. al.’s [18] the guidelines on how to conduct a scoping review. The procedures proposed by these methodologists maximize both systematicity and transparency which, in turn, ensure a high level of rigor, reliability, and trustworthiness. While scoping reviews are systematic in nature, they must not be confused with traditional systematic reviews. Indeed, whereas systematic reviews like meta-analyses attempt to integrate prior empirical findings on a mature topic in order to provide answers to questions like “what works” and “what works best,” scoping reviews attempt to provide an initial indication of the size and nature of the available literature on an emerging topic, to identify gaps, and to propose a research agenda for future work [20]. In this review, we have tried to identify the research trends in blockchain and which application domains have received most attention from the blockchain research community.

A. Volume and Trend of Research

We have surveyed three major scholarly indexing databases: Google Scholar, Scopus and Web of Science, and one reputed publishing venue (i.e., sciedencedirect). We have used “blockchain” as the keyword to search these databases. We have observed a sharp rise in volume of research over the last five years (Figure 1). If we just compare 2016 and 2017, we can see that there is 139% increase in Google scholar, 253%
increase in Scopus, 80% increase in Web of science, and a large 316% increase in sciencedirect\(^1\).

### B. Subject Areas

We have analyzed the “searched data” to find domains which have received the most attention from blockchain researchers and practitioners. We have manually gone through all the 843 entries in the scopus database. We have selected the Scopus database rather than Google Scholar because Scopus only indexes research publications whereas Google Scholar often includes non-scholarly citations. We have found 673 valid entries out of the 843 entries in Scopus. By valid entries, we mean the entries which are research publication, rather than information about the publication venues or conferences (which is often included in the search result). Figure 2 shows the numbers of the subject area-wise research publications.

![Fig. 2. Subject area wise research publication](image)

**Critical Analysis**

- Multiple parties are involved in supply chain management system and there is a trust deficit among the parties. Therefore, one entity does not reveal its information to other party.
- Currently, many practitioners solve or by-pass this problem by introducing a trusted third party. Both Party A and B trust the trusted third party and reveal their information. However, finding a suitable trusted third party is very hectic and in many cases impossible.
- Blockchain can establish trust among the transacting parties (see section V-A for trust building) without a trusted third party.

### 1) Supply Chain Provenance:

Supply Chain Management (SCM) is the management of material and information flows both in and between facilities, such as vendors, manufacturing and assembly plants and distribution centers (DC). It keeps record of different interactions between different entities until the delivery of the product to the final point of consumption. Figure 3 shows the interactions of different entities in a traditional and blockchain powered SCM. We can see that in a blockchain powered SCM, all the interactions among different parties are kept in a single blockchain ledger. Therefore, all the parties can see the transaction records related to any particular product.

Walmart uses IBM’s Hyperledger (blockchain) technology to track the source of their food and total traceability of the food at their shelf [22]. Everleger has used blockchain in diamond industry. They have used a combination of public and private blockchain to provide permissioned control and at the same time provides clear audit trail for stakeholders [23].

### 2) Inter Bank Transactions:

In current banking system, we can get real-time transactions if we transact within same bank. However, it can take 2-3 business days if the transaction is between two different banks. Situation gets worse if it is an international transaction. International transaction usually involves a third bank. The transacting banks need to have an account with this third bank, which are often referred to as Nostro accounts. The problem with current setup is time and costing.

By analyzing generic use-cases from the domains, we have tried to find out “why” and “how” blockchain is being used in these use-cases. The identified properties/criteria are discussed in more general terms in the next section.

#### 1) Supply Chain Provenance:

Supply Chain Management (SCM) is the management of material and information flows both in and between facilities, such as vendors, manufacturing and assembly plants and distribution centers (DC). It keeps record of different interactions between different entities until the delivery of the product to the final point of consumption.

Figure 3 shows the interactions of different entities in a traditional and blockchain powered SCM. We can see that in a blockchain powered SCM, all the interactions among different parties are kept in a single blockchain ledger. Therefore, all the parties can see the transaction records related to any particular product.

Walmart uses IBM’s Hyperledger (blockchain) technology to track the source of their food and total traceability of the food at their shelf [22]. Everleger has used blockchain in diamond industry. They have used a combination of public and private blockchain to provide permissioned control and at the same time provides clear audit trail for stakeholders [23].

**Critical Analysis**

- Multiple parties are involved in supply chain management system and there is a trust deficit among the parties. Therefore, one entity does not reveal its information to other party.
- Currently, many practitioners solve or by-pass this problem by introducing a trusted third party. Both Party A and B trust the trusted third party and reveal their information. However, finding a suitable trusted third party is very hectic and in many cases impossible.
- Blockchain can establish trust among the transacting parties (see section V-A for trust building) without a trusted third party.

#### 2) Inter Bank Transactions:

In current banking system, we can get real-time transactions if we transact within same bank. However, it can take 2-3 business days if the transaction is between two different banks. Situation gets worse if it is an international transaction. International transaction usually involves a third bank. The transacting banks need to have an account with this third bank, which are often referred to as Nostro accounts. The problem with current setup is time and costing.
Due the relatively higher costs of business in inter banking system, bankers are interested to see if blockchain technology can simplify and reduce the cost for inter bank payments. Some central banks such as the Monetary Authority of Singapore (MAS) and the Bank of Canada are working on solutions to use distributed ledger technology for interbank payments [29], [30]. Ripple [24], which is a crypto-currency, has been proposed to provide global settlement network based on blockchain. It has most impacted the traditional banking sector. In fact this is the first kind of crypto-currency that bridges the gap between the virtual currency market and the traditional banking sector.

**Critical Analysis**

- Blockchain has potential to be used for fast inter-bank payment transactions.
- Performance is a vital in banking system. Blockchain in its current form could not be able to handle the volume of transaction in current banking system.
- However, a consortium could be built to create and manage the crypto within that particular network. By this way the consortium can fix the price of that crypto for that network.

3) Health Care: In current health-care system, different providers keep record of their patients and they often cannot or do not share their data with other providers. Health data is very private and the patients often have to blindly trust the health care provider. In addition, accessing and sharing health record usually takes lot of admin time for the doctors and the patients. The vision of blockchain in health care system is to reduce the administrator time for the doctors so that they can spend more time with their patients and sharing of data. Researchers from MIT has proposed a blockchain based health record system called MedRec [2] which restores patients’ control over their medical data. It links patient’s medical records from different doctor’s database and allow the patient to share with any care provider.

**Critical Analysis**

- Health-care is fundamentally very complex and sensitive sector. Adaptation of technology is always very slow due to legislative requirements.
- However, interoperability and collaboration are very important in this sector for service delivery and innovation. Blockchain can be used to enable interoperability and collaboration without compromising the security of the health care providers.
- Applying blockchain in health sector without rigorous research and usability test could be catastrophic. For example, doctor’s access may get delayed due to scalability issue of blockchain in a critical moment, which may cause bad consequences.

V. **Comparison between Blockchain and Database Technology**

Table I has summarized the comparison between database and two versions of blockchain. In the following section, we will provide the details of this comparison using different criteria described in [14]. We have provided more in-depth (compared to [14]) analysis by considering different types of consensus mechanisms (in V-D), and attack vectors (e.g., DDos (in V-C), 51% attack (in V-E)).

A. **Trust Building**

One of the most important features of blockchain technology is immutability. The immutability is achieved by decentralized consensus mechanism. Each participating node takes part in a consensus mechanism to check whether any particular transaction is valid or not. Every node in the system has the same level of access (e.g., in terms of public blockchain) and capability. This provides a solid foundation for building trust, because it democratize the whole system. In a traditional database, we have to rely on a single central authority which controls who can do what in the system. This kind of system is good when the party who controls the system is trusted and behaves honestly.

B. **Confidentiality and Privacy**

There is a misconception about blockchain is that data in blockchain is kept encrypted. However, this is not true. The data is digitally signed by the the transacting parties but not by default encrypted. In fact, it is an open ledger system, where anybody can join and verify any transaction in the network. However, the privacy or confidentiality of the participating parties are kept by using public key cryptography. The transactions reveal the transacting parties and the data in the transaction (e.g., amount of coin in case of crypto-currency). Recently, researchers are proposing strong anonymization using cryptographic means such as Zero Knowledge protocol [31].

C. **Robustness/Fault Tolerance**

Blockchain is a decentralized system and uses distributed computing mechanism to provide robustness and fault tolerance. Data in the blockchain is stored distributedly. Each participating nodes stores a copy of the blockchain. Therefore, it can all types of cyber attacker that is relevant for single point of failure. Attacks like Denial of Service (DoS) and Distributed Denial of Service (DDoS) are infeasible in blockchain network. If any particular node goes down or got compromised then other nodes can still continue the job.

D. **Performance**

Blockchain, specially bitcoin is notoriously slow. It takes 10 minutes to confirm a transaction in the network. This time could go up to 60 minutes if there is any soft fork [26] happens in the network. Traditional system database systems can be designed to handle thousands of transactions per second. For example Visa and Mastercard networks can handle 50,000 transactions per second. If the system administrator finds a performance bottle neck, he can replace or re-engineer the system to allow high volume of transaction.

However, in terms of performance there are research going on to improve the efficiency of the consensus mechanism.
Consensus algorithm like Ethash [27] and X13 [28] can come to an consensus within 10 to 20 seconds.

### E. Security

The security of blockchain comes, in part, from its adaptability. The more users the system has the more users can be required to achieve consensus. In Blockchain protocol, a block will be accepted if 51% of the mining nodes agree. Therefore, if 51% of the mining nodes are controlled by the malicious users then an “invalid transaction” can be accepted as “valid transactions”. If there are enough people in the network, it seems impossible but it can happen.

In terms of traditional database, the state of the database is maintained by a central system. Access to the data is restricted by the access control mechanism set by that system. This system is vulnerable, if the system administrator of the system is compromised.

### VI. WHEN SHOULD WE USE BLOCKCHAIN?

From the above discussion, it should be clear that blockchain is not a general purpose technology but should be applied judiciously to reap its benefits. Figure 4 provides a simple decision flow control diagram that can be used by the business analyst or system architect to decide whether blockchain should be used or not.

In general, blockchain technology is useful in those use-cases where there are more than one administrative authority and there is a trust deficit among those parties. A typical example could be a supply chain management system, where multiple parties collaborate together to deliver goods. Another example could be a consortium of independent companies working on a government project, where there is a trust deficit between the parties.

Currently, if there is a trust deficit among the collaborative parties, they usually select a trusted third party which both parties trust. For example, both buyer and seller trust bank to transact money among themselves. However, sometimes finding a trusted third party is challenging or risk prone.

If all the criteria above match then we can say that blockchain is a beneficial technology for that use-case. However, now we have to decide whether we should use public blockchain or private blockchain. If the stored value in the blockchain needs to be publicly verifiable then we should use public blockchain, whereas if the data is only for specific parties then a private blockchain is a better choice because consensus mechanisms can be simplified.

### VII. ASSESSMENT OF EXISTING USE OF BLOCKCHAIN

We have analyzed a sample of 100 papers among the 800+ papers surveyed from Scopus using our decision tree model (see figure 4). We have found that the majority of the 100 papers are theoretical papers related to core blockchain technology. We have identified 34 papers which are related to use cases or their implementations. In our analysis, we have found that among these 34 papers, 25 papers have used the blockchain appropriately (e.g., used the unique properties of blockchain), whereas the used cases in the other 9 papers could have used traditional databases to achieve the same functionality.

We have found that 9 papers out of 25 papers (which used the unique properties of blockchain) are related to supply chain
management system and power grid. Blockchain has emerged as a productive technology for supply chain management (see section IV-1). Home-based renewable energy generation and distribution management system is another promising area for the use of blockchain.

An example of inappropriate use of blockchain can be smart home scenarios. Firstly, smart home is a private network and usually does not have any trust issues. The immutable and decentralization nature of the blockchain is also not very relevant for smart home scenarios.

The list of papers and their short analysis and decision is given in an online spreadsheet and can be found at https://tinyurl.com/yb3pvhsa.

VIII. CONCLUSION

Blockchain has gained much attention from both researchers and practitioners. They have tried to use it in solving different kinds of problems. In this paper, through a scoping survey, we have identified areas that have received the most attention from the blockchain research community. Then we have carried out a critical analysis of these domains, regarding their need for blockchain. We have also done a critical comparative analysis for blockchain and traditional database system in terms of the range of properties used to evaluate any information system. We have found that if trust building, robustness, and provenance of data are the priorities of the system, then blockchain is the better solution. If confidentiality and performance are the main concerns, then traditional database is still the better solution.

Finally, we have provided a decision tree to evaluate the appropriate use of blockchain (vs. database technology) depending on the properties of the problem with the aim of helping to avoid mis-uses of blockchain and unsuccessful system implementations.

REFERENCES

[8] Sanamiego M, Deters R. Blockchain as a Service for IoT. InInternet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2016 IEEE International Conference on 2016 Dec 15 (pp. 433-436). IEEE.