Mindanao Journal of Science and Technology Vol. 19 (1) (2021) 137-162

Blockchain-based Big Data Integrity Service Framework for IoT Devices Data Processing in Smart Cities

Tanweer Alam
Faculty of Computer and Information Systems
Islamic University of Madinah
Madinah, 42351 Saudi Arabia
tanweer03@iu.edu.sa

Date received: July 7, 2020
Revision accepted: January 28, 2021

Abstract

Nowadays, smart cities are using advanced technologies to control and coordinate physical, social, and commercial enterprise systems to deliver high-quality services to their residents while guaranteeing the effective usage of available resources. Numerous major corporations, as well as government bodies, are involved in the construction of smart cities because smart city activities also need information that is obtained from organizations. Blockchain-based technology allows for peer-to-peer exchanges as a public ledger that maintains records of transactions, smart contracts, agreements, and shipments without third-party mediators. In the future, smart cities will be equipped with several innovative technologies, a huge amount of data, a lot of Internet of Things (IoT) devices, and a heterogeneous environment. However, data processing among IoT devices in futuristic smart cities is a challenging task. In this paper, the author proposed a blockchain-based big data integrity service framework for IoT devices data processing in smart cities. The author collected the necessary data from various resources and applied three experiments to evaluate the key performance indicators. K-mean algorithm was used to classify the data, and blockchain strategy was applied to ensure secure communication among IoT devices. The framework was simulated and tested using 100 devices. The proposed framework allowed IoT devices to efficiently use information collected from different resources to engage in operational processes and exchange information.

Keywords: big data, blockchain, smart cities, internet of things, secure communication

1. Introduction

The vast amount of research on big data analytics, blockchain, and the integration of both technologies provides many possibilities for smart city applications to thrive using these technologies (Karafiloski and Mishev, 2017). The acceptance of blockchains by companies across the globe keeps growing due to its success and the latest big data innovations have made data from different sources accessible besides investigation (Alam *et al*., 2020).

Throughout the world, several cities are trying to become smart cities. Even so, its expertise and the system on its use of information by smart cities are still largely undefined (Alam and Benaida, 2018). Blockchain and big data growth have played a key role in the efficacy of smart city developments (Hassani *et al*., 2019). However, a pertinent question remains: Are the current technologies capable of achieving a security solution inside the smart cities that are massively establishing?

Various steps should be considered, besides securing only sustainable development (Alam, 2020). Internet-connected gadgets like sensors, smart devices, and cameras making up the Internet of Things (IoT) keep rising exponentially (Alam, 2017). The trends of big data, blockchain, and smart cities from January 2016 to January 2021, according to Google Trends, are shown in Figure 1 (Google Trends, n.d.).

The objective of this study was to create a novel big-data- and blockchain-based integrated framework for IoT device data processing for futuristic smart cities. In this article, the author discussed such an integrated approach to assessing data exchange in smart cities, which helps to share resources across different unauthenticated environments. This integrated technology allows smart devices to efficiently use data collected from different resources to engage in operational processes without requiring access to the main information.

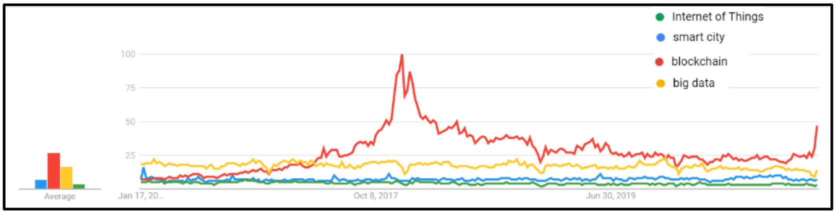


Figure 1. Big data, blockchain, IOT, and smart city trends from 2016 to 2021 worldwide

A recent International Data Corporation (IDC) study predicted that there could be 41.6 billion IoT-based smart machines, or smart things, by 2025, which could produce 79.4 zettabytes of information (Shirer and MacGillivray, 2019). The internet of smart devices is rising exponentially day by day. The number of IoT-connected devices is expected to reach around 22 billion by 2025 (Figure 2) (Lueth, 2018). Blockchain technology could become incredibly profitable by 2030, making up around 20% of all big data business and generating approximately $100 billion in annual revenue. This will translate to over $100 billion in sales yearly (Fedak, 2018).

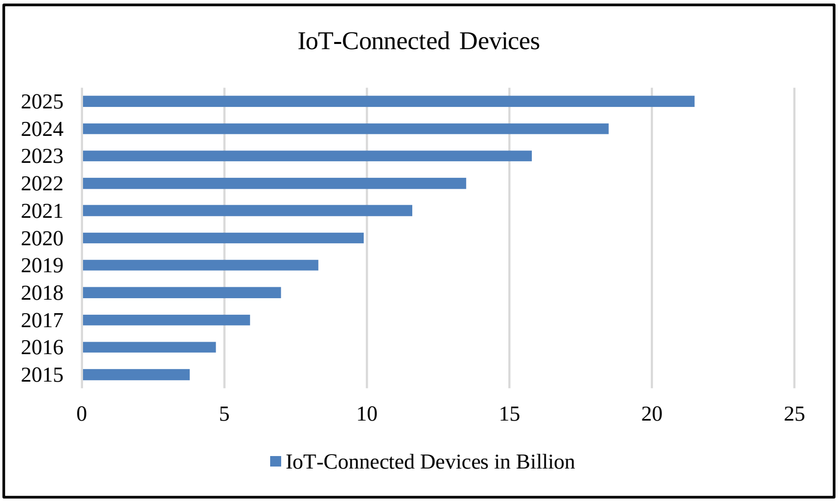


Figure 2. IoT connected devices from 2015 to 2025

Blockchain technology is intended to act as a security measure for decentralized networks. Transactional data and records are circulated, exchanged, and used to establish a universal repository with all transmissions among IoT devices (Alam, 2019a). This research is as something of a novel of factors, just in a public, common, and standardized effort to create consensus and confidence, without all of the mediators of third-party service providers, in a direct interaction among concerned IoT devices. This grows steadily as a new batch of transactions introduces additional full blocks. All of the blocks are directly stored on the blockchain without using a third party. Big data analytics is a procedure by which high quantities of records are examined and only helpful data are obtained for further investigation (Chen *et al.*, 2019).

Currently, there is no specific method for information collection in smart cities. The information can be organized, unorganized, or semi-organized. Information collection, preservation, processing, scanning, transmission, uploading, accessing, compiling, modification, and protection, as well as the database, are major problems in big data analytics.

Big data is related to three essential aspects: quantity, wide range, and time. Users cannot test big data by themselves, but they can read and monitor how it occurs because users can process large amounts of information. Consequently, big data mostly involves information that surpasses what conventional technology is able to process at a reasonable speed and cost. Blockchain has become a mechanism of storing information to validate and verify exchanges by consensus between all parties concerned (Alam, 2019b). The architecture is based on the idea of big data processing and mining that tracks occurrences. Smart cities utilize infrastructure to enhance information facilities and people's capabilities.

Throughout the processing of specific information, such as traffic flow, electricity consumption, and air quality, smart cities can utilise IoT tools, systems, and applications (Alam and Benaida, 2019). Figure 3 shows the role of big data and blockchain in smart cities. In this study, the contributions of the author are to explore the integration of big data and blockchain technologies and create a framework in which IoT devices can exchange processed and valuable data securely. Big data analytics provides valuable data from the large amount of data stored in big data, and blockchain provides secure communication among IoT nodes.

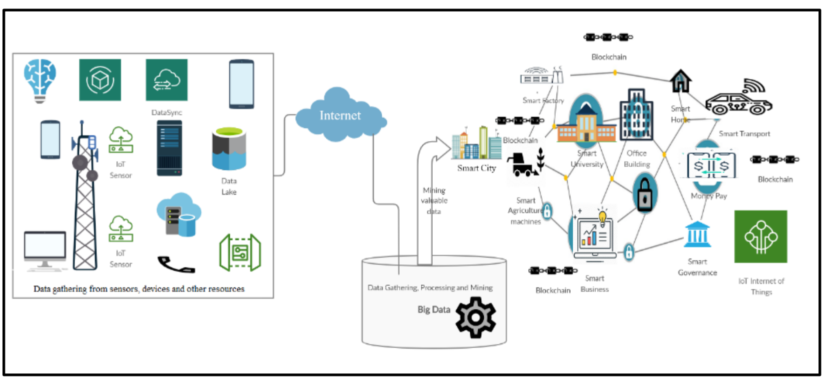


Figure 3. Blockchain-based big data analytics framework for smart cities

In this study, the author gathered data from various resources, stored them in big data, and used the dataset of the big data. After data gathering, the key performance indicators were applied and their values were obtained. Application of big data analytics to mine the data and find valuable information was then carried out after applying the key performance indicators. Following the extraction of valuable data, blockchain technology for secure data transmission among the IoT nodes was then employed.

In smart cities, blockchain and big data play essential roles by continuously rendering available information and gathered aspects of data processing. This process functions through interconnection among linked sensors and data sharing, which involves the internet, wireless connections, and other media. In general, smart cities utilize IoT tools to capture and easily process information in testing areas (Alam and Aljohani, 2015a). Wireless infrastructure devices and embedded apps receive and evaluate data to make smarter decisions based on information collected from existing resources.

1.1 Advantages of a Smart City

1.1.1 Effective Use of Resources

It is essential to incorporate ideas to improve the monitored use of resource management with several services, whether they are unavailable or highly complicated. Through continuing management programs, it becomes easier to effectively identify and disperse services, thus minimizing the cost or increasing the use of electricity and resources. Therefore, one essential feature of smart city systems is the architecture for interconnection and data processing, which often makes for greater communication between apps and utilities.

1.1.2 Best Standard of Living

A good standard of living for smart city residents covers improved services and quality of living. It is the product of increased preparation of lifestyle/workplaces or sites, improved and simpler infrastructures, and the provision of adequate knowledge of intelligent decision-making.

1.1.3 Increased Standard of Visibility and Flexibility

The desire to properly monitor and regulate the features, characteristics, and technologies of a smart city would improve interoperability and accessibility. The norm involves sharing information and distributing data. An efficient smart city would, therefore, increase the accessibility of information for those concerned. It would encourage cooperation and coordination between organizations and build smarter city-enhancing infrastructure and apps.

An example is a US administration that, through the pursuit of transparency and accountability, has gathered and published a wide variety of documents, articles, and content. Smart energy, smart logistics, smart universities, smart homes, smart hospitals, smart security, smart buildings, smart transport, smart learning, smart vehicles, smart hospitality, smart payments, smart business, and smart governance are applications of smart cities (Figure 4).

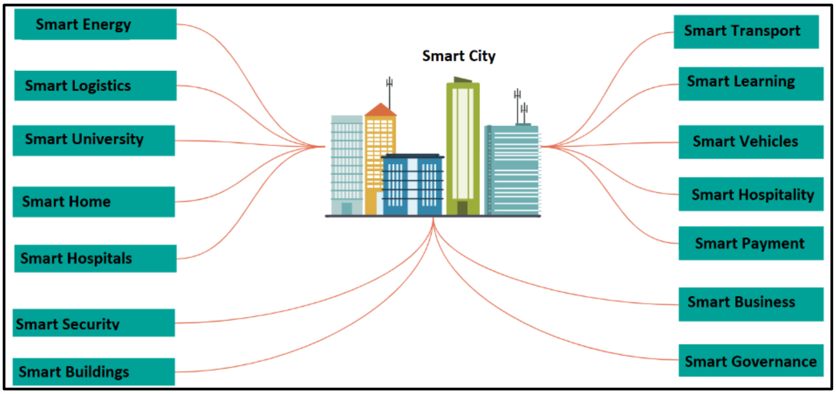


Figure 4. Smart city applications

The most well-known example of smart city implementation is seen in Dubai, where the government is exploiting numerous technological advances, such as blockchain, big data, and intelligent systems, to make Dubai "the happiest city on Earth" and the very first city completely driven by blockchain. Almost a year earlier, the Crown Prince of Dubai revealed that Dubai would be the first blockchain-based smart city to secure government records by 2021. The blockchain policy of Dubai is based on three factors: productivity of states, business growth, and foreign strategic vision. The Smart Dubai Office (SDO) introduced the Blockchain Project very quickly to enable start-ups to drive blockchain initiatives. Throughout the Emirates, this development has the potential to redefine banking, economic growth, foreign direct money and transfers among others (Bishr, 2019).

1.2 Previous Studies

The broad data study for IoT and smart cities was discussed by Strohbach *et al*. (2011). The findings can be summarized in this article. First, an example in the smart environment was introduced to demonstrate the connection between broad-scale information and IoT technology, as well as the high-level needs for these large computational databases (Strohbach *et al*., 2015).

Big data and smart urban planning have been highlighted and describe smart cities equipped with intelligent technologies and networks that utilize big data based on several situations (Kitchin, 2014). This technology enables smart city campaigners to track living environments in more secure, efficient, safe, serviced, inexpensive, and sensitive neighborhoods in actual environments and provides various modes of public management (Kitchin, 2014).

The concept of IoT and big data for smart cities was introduced by Sun *et al.* (2016). Innovative and interconnected cities originated from the vision of digital technology. The aim was to address the need for culture (sustainability and renovation). These were introduced to fulfil a need to survive in today's world (standard of living) and to plan for the long term (sustainability). This is also believed to boost living standards, recycling, regeneration, and balanced urban growth. Some of the goals of developing an intelligent community are to survive in the present, build a better future, and look at the past (Sun *et al*., 2016).

The design of the smart network is described by Alfa *et al*. (2018) as an effective connectivity mechanism with other networks. Increased performance of every device would be needed to make optimal use of usable resources at low latency for rapid decision-making processes (Alfa *et al*., 2018).

Big data analytics has been addressed relating to smart governments in future smart cities, and the topic of efficient governments has been addressed too. The subject was one of the fields in which the various facets of how smart governance could be translated into successful governance were studied. Big data analytics has been applied within businesses and organizations by increasing the visibility of their business problems through profitability, and by delivering support teams. Big data analytics acts as a tool in extending its current applications (Kulkarni and Akhilesh, 2020).

1.3 Role of Big Data and Blockchain in Smart Cities

Nowadays, the everyday lives of people are set under a digital umbrella, including their ability to think, real-time social activities, shopping practices, information searching, photographs, and videos. Around the world, most types of decision-makers, from business leaders to public officials to scientists and researchers, want to focus their decisions and choices on information. In order to find relevant information, a modern big data analytics approach is developed. Big data analytics is an approach for distilling a large amount of information down to a specific bit of highly valuable information.

Blockchain technology is a shared database-based type of technology. A blockchain network is a shared transaction repository between different networks. An action is loaded in the repository and is recorded as a block in the chain. A blockchain is a chain among which such blocks have been formed. A blockchain system would be a permanent centralized knowledge base, wherein a single time-stamped operation may be inserted and clustered into a chain of keys. The specification describes many versions of these blocks that can be stored in and built into a shared database. Through opportunities for people and businesses, the advancement of big data and blockchains is increasingly expanding and improving most fields of innovation. The volume of information generated through IoT devices will play a significant role in the world of big data. Big data is classified by three components: length, range, and speed. The ability to analyze and leverage tremendous quantities of IoT data, such as smart cities, smart transportation, energy consumption, and smart grids, offers significant potential.

In smart cities, a large number of connected devices are equipped in IoT networks using sensor networks. Different systems receive information from different sensors and transmit it using a connected sensor network. A range of connectivity technologies, such as Bluetooth, Wi-Fi, WSN, and wireless communications, interlink connected devices and machines. The sensor nodes send and receive information from fully autonomous systems that enable better compatibility with smart devices to improve the quality of life of the residents of smart cities.

2. Methodology

Examining large amounts of information and mining valuable information are not easy tasks. However, data have been collected and stored for a long time. Information gathering is generally simpler than extracting the required information from a large amount of information. Although modern computing technologies have become quicker than the computing of the 19^{th} century, collecting information on a grand scale is a challenge for the machines that are needed for smart cities today. Several successful strategies, such as sampling, information condensation, density-based framework, grid computing, reinforcement techniques, and cloud services, have been proposed as solutions to criticisms of the analysis of a large amount of information (Tsai *et al*., 2015).

Information that needs to be examined is abundant and gathered from different sources of IoT. However, it is made up of different kinds of information, such as data collected from sensors, IoT devices, or another connected network. When the number of sensors and smart devices increases, a large amount of information is generated by such machines. A small amount of that information is brief or unusual, providing a reasonable statistic of a device's fitness, whereas visual surveillance systems can produce massive amounts of information through artificial intelligence, for instance, to monitor massive crowds. It can depend on the connectivity of the nodes and the information gathered by the IoT devices. Big data analytics could enhance strategies for data mining and data observation. Big data allows more information to be analyzed to discover further relevant data, and more data does not automatically equal more valuable knowledge. This could lead to some vague and anomalous results.

The K-mean clustering approach, which was used in this study, is a type of unsupervised machine learning that is able to classify unorganized information beyond specified clusters. This algorithm involves selecting categories within the information, including the set of organizations where the parameter k signifies. The algorithm assigned every piece of information to one of the k clusters based on specified characteristics. The K-mean cluster relies upon this concept of range for allocating datasets to the clusters within the process.

The K-mean approach is a classification technique involved in splitting data into different groups of a sequence of n objects. The pieces in a group would be similar and are separated from others in some other groups. Assume $\mathscr{R}$ is a collection of n things separated into several groups $\mathcal{Z}$. This sums up the technique as described below.

Step 1: Initialize $\mathscr{R}$ and $\mathcal{Z}$ as follows.

$$\mathscr{R} = \{\mathscr{R}_1, \mathscr{R}_2, \ldots\ldots \mathscr{R}n\}$$
$$\mathcal{Z} = (\mathcal{Z}_1, \mathcal{Z}_2, \ldots\ldots \mathcal{Z}k)$$

Step 2: Classify. Get the Euclidian distance from each $\mathscr{R}_i$ to k centroid.

Step 3: Assign each $\mathscr{R}_i$ to the nearest $\mathcal{Z}_i$. Find $\mathcal{Z}_i$.

Step 4: Convergence. If $\mathcal{Z} = (\mathcal{Z}_1, \mathcal{Z}_2, \ldots\ldots \mathcal{Z}_k)$ remains unchanged for twice consecutively, stop; otherwise, go to step 2.

Consider the measurement of data that this will accomplish. Below are essential characteristics, described in the ranking of differentiation: popular properties are a, b, and c, but d, e, and f are used sometimes (Granville, 2013).

a. The variation will be zero when the variations $v_1=v_2= \ldots =v_n$ values.

b. A variation becomes symmetry. Each $(v_1, \ldots, v_n)$ possible combination has the same variation to $V(v_1, \ldots, v_n)$.

c. In any fixed value f, the variation is called fixed-invariant: $V(v_1+f, \ldots v_n+f) = V(v_1, \ldots v_n)$.

d. A variation becomes restricting if $V(v_1, \ldots v_n) < 1$.

e. If the fixed value is 1 then the variation for every level value $V(f*v_1, \ldots, f*v_n) = V(v_1, \ldots, v_n)$.

f. A variation through a variety of vector functions like motions, or projections, is unobservable.

Theorem: The equation $V(v_1, \ldots v_n) = \{\sum_{i=1}^{n} vi/n\}-c(\sum_{i=1}^{n} vi/n)$ describes a variation that satisfies a and b conditions, while c is a strongly convex function.

The problem within Hadoop and Maps Reducer is the time spent transmitting the information when it divides the job into many executive operations that execute on different machines. All of these factors shift the computation efficiency issue from the complexity of the algorithms to the spent of information exchange or processing power. Interaction implies the duration in locating inside from a cloud platform or computing resources to memory or storage devices. The issues are as follows: Why would such transmissions be optimized? Why will simulations in storage be optimized? Users cannot depend on an efficient power like variation since no such powers are executed completely.

Blockchain is a series of blocks (Bruyn, 2017). The size of a block (*T*) in a blockchain can be obtained by using Equation 1.

$$T\ (in\ MB) = H\ (in\ bytes) + S\ (in\ bytes) \times No\ of\ transactions\ in\ that\ block \tag{1}$$

where:

H = size of the header
S = size of transaction

Total mining energy (*Em*) can be calculated by the sum of all connected IoT nodes mining energy (*Et*) (Equation 2).

$$Em = \sum_{i=1}^{n} Et(i) \tag{2}$$

Time (*ti*) is taken by a block between interconnected IoT nodes with their bandwidth (*Bn*) (Equation 3).

$$ti = T / Bn(i,j) \tag{3}$$

where:

$Bn(i,j)$ = bandwidth between the IoT nodes i,j.

The framework width (*W*) can be obtained using Equation 4.

$$W = Max\ (ti) \tag{4}$$

The hash is the crypto code of the previous block (*Hc*) assigned to every block in the blockchain. The block occurrence (*Bo*) can be derived through Equation 5.

$$Bo = Te/Hc \tag{5}$$

The transactions per second (*Ts*) held in the blockchain can be calculated using Equation 6.

$$Ts = Bo \times Number\ of\ transactions\ in\ that\ block \tag{6}$$

A blockchain can fork if the miners received multiple blocks at the same time. The probability of a blockchain fork *P(F)* occurring can be calculated by following Equation 7.

$$P(F) = 1 - (1+Bo\times Fw) \times e - Bo\times Fw \tag{7}$$

Consider the newly created block *BN* that is referenced by two previously created blocks. The *BN* is placed between these two blocks. Consider that the block *BN* is verified by the miners; the time *BN(t)* is represented as the verification time calculated by the Poisson process μ.

Suppose that α is the time taken by the IoT node to compute the transaction process to the fog network. Therefore, the total time taken by each transaction to be visible on the network is α+t [9]. Consider that *N* is the total number of unverified blocks. The probability of *BN* in the blockchain and big data is obtained using Equation 8.

$$P(BN) = N / (N + \mu\alpha) \tag{8}$$

For example, consider the 5 IoT nodes connected with the P2P network with the mining energy 500, 1000, 300, 700, and 1500, respectively. They can do transactions in the proposed framework. Suppose that the size of the header is 100 bytes and the size of the transaction is 500 bytes. The IoT nodes in the framework are five.

Using Equation 2, the block size is calculated.

$$BS = 100 + 500 \times 5 = 2600\ MB$$

By using Equation 2, the total mining energy is derived.

$$Em = \textstyle\sum_{i=1}^{n} Et(i)$$

$$= 500 + 1000 + 300 + 700 + 1500$$
$$= 4000$$

Suppose BS = 1, MB = 1,048,576 bytes. Calculate the total number of transactions in a block = (BS – HS)/TS.

$$= 1{,}048{,}576\text{-}100/500$$
$$= 2096.952$$

Eventually, this research has generated new equation – Equation 8 – to better clarify improved big data analytics with blockchains for smart cities.

Main performance metrics such as mean absolute deviation (MAD), mean squared error (MSE), root mean squared error (RMSE), and mean absolute percentage error (MAPE) can be observed in the experiments in the succeeding section. MAD is the mean variance among the sample points and the mean that gives the sense of the uncertainty in the information gathering. The variability of the estimation techniques is called MSE. RMSE is a commonly applied indicator of the variations among the results expected by the system or measurement device and the observed values. MAPE is an indicator of the precision of the forecasting processes. The formulas to obtain MAD, MSE, RMSE, and MAPE are as follows.

$$MAD = \frac{\sum_{t=1}^{n} |A_t - F_t|}{n} \tag{9}$$

$$MSE = \frac{\sum_{t=1}^{n} (A_t - F_t)^2}{n} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (A_t - F_t)^2}{n}} \tag{11}$$

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|}{n} \times 100 \tag{12}$$

where:

A = actual value
F = forecast value

3. Results and Discussion

Big data analytics enables the mechanisms for the seeking, extracting, and assessing of information committed to enhancing the efficiency of data transmission in smart cities. It is an approach to extract valuable information from a large amount of information. A blockchain is a list of cryptographically secure blocks of valuable information (Alam & Aljohani, 2015b). Blockchains demonstrate that confidential activities should always be permitted. Networks are extremely vulnerable, but blockchains can use them to do secure transactions. Blockchains can solve the issue of confidentiality in the lines of trusted responsibility. Multiple smart devices are connected through wireless links to each other, and they become smarter as they inform each other that public networks have become vulnerable. Blockchains provide solutions by means of sequential and recently discovered information that could be generated.

3.1 Setup for Experiments

3.1.1 Datasets

A dataset is a collection of input data received from different resources, such as sensors (different kinds or purposes), IoT devices (smart gadgets), or physical things. A dataset has multiple files, including sensors, traffic, road, parking and pollution among others. The dataset used in this article is an expanded version of the “Multi-source dataset for urban computing in a Smart City” (Honarvar and Sami, 2019), which has been applied to various resources from multiple input resources.

3.1.2 Big Data Analytics

Big data analytics is a technique that uses efficient scientific methods to process, mine, and obtain valuable information from a large amount of information (either organized, semi-organized, or unorganized) from multiple resources. Hadoop was used in the experiments.

3.1.3 Blockchain

A blockchain is an excessive collection of records, known as blocks that become connected through cryptography. Every block consists of a sequence number including its transaction’s hashing algorithm and transaction information. After processing, the relevant information will be transmitted to the blockchain. The blockchain will then create a block of information and send it to the network. All departments of smart cities are connected in the blockchain network through which the information will be transmitted to the relevant department. In this study, Hyperledger was utilized for blockchain implementation.

3.2 Analysis of Results

Big data analytics processes a lot of data for different purposes. This study focused on traffic data received from multiple traffic sensors. MAD, MSE, RMSE, and MAPE were obtained for the time intervals of 5, 10, 15, …. 60 min. Figures 5, 6, and 7 represent the MAD, MSE, RMSE, and MAPE values in different experiments on the dataset. MAD is the sum of absolute differences between the actual value and the forecast divided by the number of observations. MSE is the sum of the squared errors divided by the number of observations; RMSE is the square root of the MSE; and MAPE is the average of absolute errors divided by actual observation values. Tables 1, 2, and 3 show the three experiments that were carried out to obtain MAD, MSE, RMSE, and MAPE values.

Blockchain networks need miners for different testing purposes. Initially, two miners were nominated with 10 transactions and set the block demand for each miner. Next, five miners with 50 transactions were nominated and set the demands for each miner. When the demands increase, the possibility that the miner mines the block also increase. Computational requirements such as processor utilization, memory management in blockchain, and big data were evaluated. The findings showed that the processor used for blockchain was lower than that for the big data. Memory usage varies based on the number of blocks and transactions. When the number of blocks in the chain increases, then memory usage also varies. Likewise, if the number of transactions varies, then the memory usage also varies. Figure 8 shows the memory and processor usage in blockchain and big data. Figure 9 shows the transmission delay computation while Figure 10 discloses the memory usage and number of blocks versus transactions in blockchains.

Figures 11 to 14 shows the performance evaluation of blockchains with 5, 10, 50, and 100 smart devices. Blockchains are aimed at decentralization as a security measure. Blocks may transmit and communicate information or files, which make a massive measure with all records occurring in a sector of the smart city. This is also steadily increasing as a bunch of new files bring in new full units. Now, a group of 47 Japanese banking institutions are collaborating with a blockchain group named Ripple to allow cash transactions through blockchains between bank account holders. The primary goal of this initiative is to perform real-time transactions at a considerably affordable price. The Smart Dubai Project is a public-private collaboration to create a sustainable environment via an outstanding service for all of its residents, as well as an effective ledger-based, technology-driven city administration.

Table 1. Obtained MAD, MSE, RMSE, and MAPE values (Experiment 1)

Time Duration	Observation Data	Simulation or Output Data	Error	Absolute Value of Error	Square of Error	Absolute Values of Errors Divided by Actual Values
i	Oi	Si (Theta1)	Oi-Si	| Oi -Si|	(Oi-Si)^2	| (Oi-Si)/Oi|
5	20746220.000	158324.000	20587896.000	20587896.000	423861461706816.000	0.9924
10	20746392.000	158324.000	20588068.000	20588068.000	423868543972624.000	0.9924
15	20746723.000	158324.000	20588399.000	20588399.000	423882173383201.000	0.9924
20	20747172.000	158324.000	20588848.000	20588848.000	423900661967104.000	0.9924
25	20747545.000	158324.000	20589221.000	20589221.000	423916021386841.000	0.9924
30	20747994.000	158324.000	20589670.000	20589670.000	423934510708900.000	0.9924
35	20748443.000	158324.000	20590119.000	20590119.000	423953000434161.000	0.9924
40	20748892.000	158324.000	20590568.000	20590568.000	423971490562624.000	0.9924
45	20749341.000	158324.000	20591017.000	20591017.000	423989981094289.000	0.9924
50	20749790.000	158324.000	20591466.000	20591466.000	424008472029156.000	0.9924
55	20750239.000	158324.000	20591915.000	20591915.000	424026963367225.000	0.9924
60	20750688.000	158324.000	20592364.000	20592364.000	424045455108496.000	0.9924
	Totals		205895272.0	205895272.0	4239286317245720.0	9.924
	n	12				
	MAD	17157939.333				
	MSE	353273859770476.000				
	RMSE	18795580.857				
	MAPE	82.70				

Table 2. Obtained MAD, MSE, RMSE, and MAPE values (Experiment 2)

Time Duration	Observation Data	Simulation or Output Data	Error	Absolute Value of Error	Square of Error	Absolute Values of Errors Divided by Actual Values
i	Oi	Si (Theta2)	Oi-Si	| Oi -Si|	(Oi-Si)^2	| (Oi-Si)/Oi|
5	20752484.000	158324.000	20594160.000	20594160.000	424119426105600.000	0.9924
10	20752933.000	158324.000	20594609.000	20594609.000	424137919862881.000	0.9924
15	20753382.000	158324.000	20595058.000	20595058.000	424156414023364.000	0.9924
20	20753831.000	158324.000	20595507.000	20595507.000	424174908587049.000	0.9924
25	20754280.000	158324.000	20595956.000	20595956.000	424193403553936.000	0.9924
30	20754729.000	158324.000	20596405.000	20596405.000	424211898924025.000	0.9924
35	20755178.000	158324.000	20596854.000	20596854.000	424230394697316.000	0.9924
40	20755627.000	158324.000	20597303.000	20597303.000	424248890873809.000	0.9924
45	20756076.000	158324.000	20597752.000	20597752.000	424267387453504.000	0.9924
50	20756525.000	158324.000	20598201.000	20598201.000	424285884436401.000	0.9924
55	20756974.000	158324.000	20598650.000	20598650.000	424304381822500.000	0.9924
60	20757423.000	158324.000	20599099.000	20599099.000	424322879611801.000	0.9924
	Totals		205961805.00	205961805.00	4242026528517880.000	9.924
	n	12				
	MAD	17163483.750				
	MSE	353502210709824.000				
	RMSE	18801654.467				
	MAPE	82.70				

Table 3. Obtained MAD, MSE, RMSE, and MAPE values (Experiment 3)

Time Duration	Observation Data	Simulation or Output Data	Error	Absolute Value of Error	Square of Error	Absolute Values of Errors Divided by Actual Values
i	Oi	Si (Theta3)	Oi-Si	| Oi -Si|	(Oi-Si)^2	| (Oi-Si)/Oi|
5	20763260.000	158324.000	20604936.000	20604936.000	424563387564096.000	0.9924
10	20763709.000	158324.000	20605385.000	20605385.000	424581890998225.000	0.9924
15	20764158.000	158324.000	20605834.000	20605834.000	424600394835556.000	0.9924
20	20764606.000	158324.000	20606282.000	20606282.000	424618857863524.000	0.9924
25	20765055.000	158324.000	20606731.000	20606731.000	424637362506361.000	0.9924
30	20765504.000	158324.000	20607180.000	20607180.000	424655867552400.000	0.9924
35	20765953.000	158324.000	20607629.000	20607629.000	424674373001641.000	0.9924
40	20766402.000	158324.000	20608078.000	20608078.000	424692878854084.000	0.9924
45	20766851.000	158324.000	20608527.000	20608527.000	424711385109729.000	0.9924
50	20767300.000	158324.000	20608976.000	20608976.000	424729891768576.000	0.9924
55	20767749.000	158324.000	20609425.000	20609425.000	424748398830625.000	0.9924
60	20768198.000	158324.000	20609874.000	20609874.000	424766906295876.000	0.9924
	Totals		206069558.000	206069558.000	4246466290054190.000	9.924
	n	12				
	MAD	17172463.167				
	MSE	353872190837849.000				
	RMSE	18811490.925				
	MAPE	82.70				

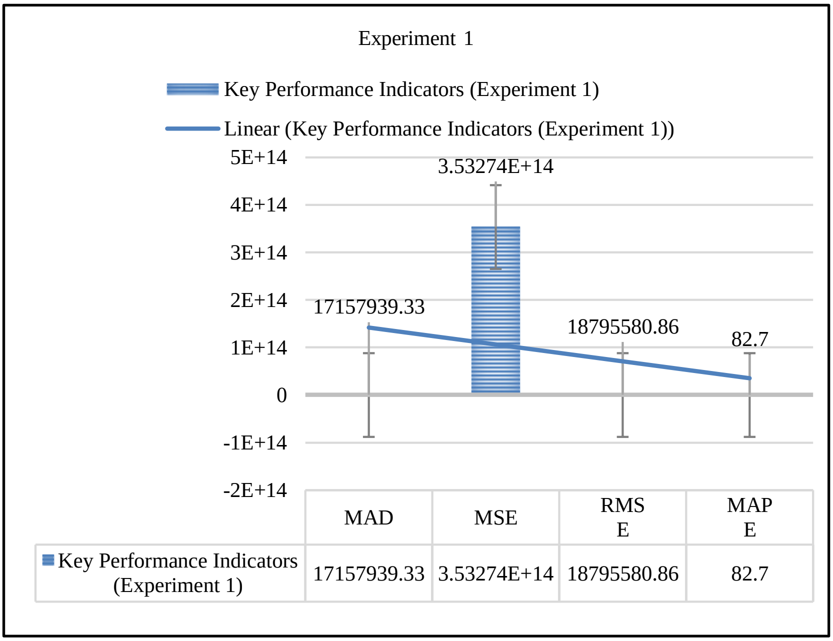


Figure 5. The values of MAD, MSE, RMSE, and MAPE in Experiment 1

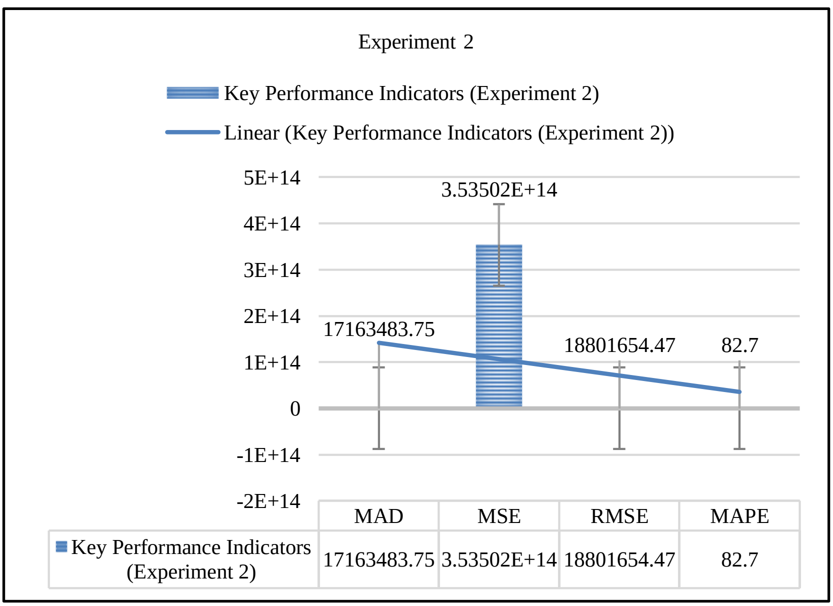


Figure 6. The values of MAD, MSE, RMSE, and MAPE in Experiment 2

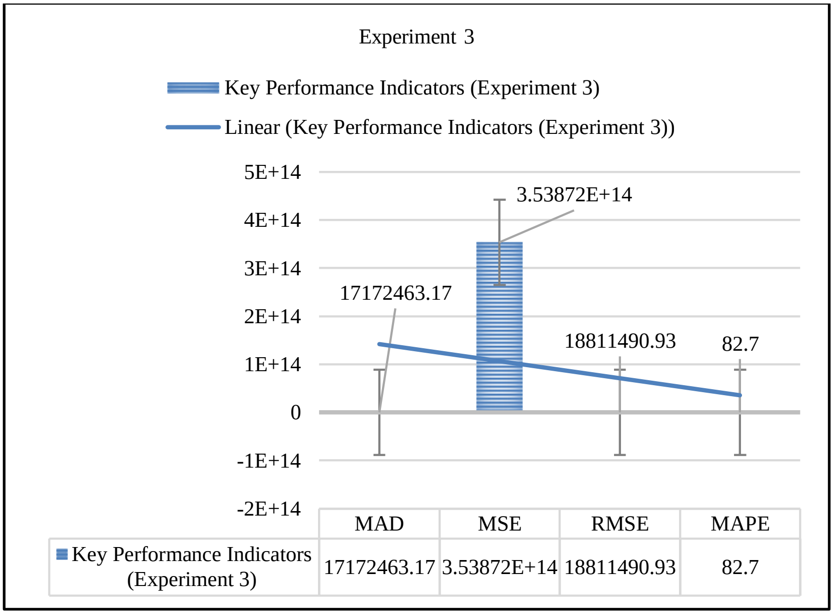


Figure 7. The values of MAD, MSE, RMSE, and MAPE in Experiment 3

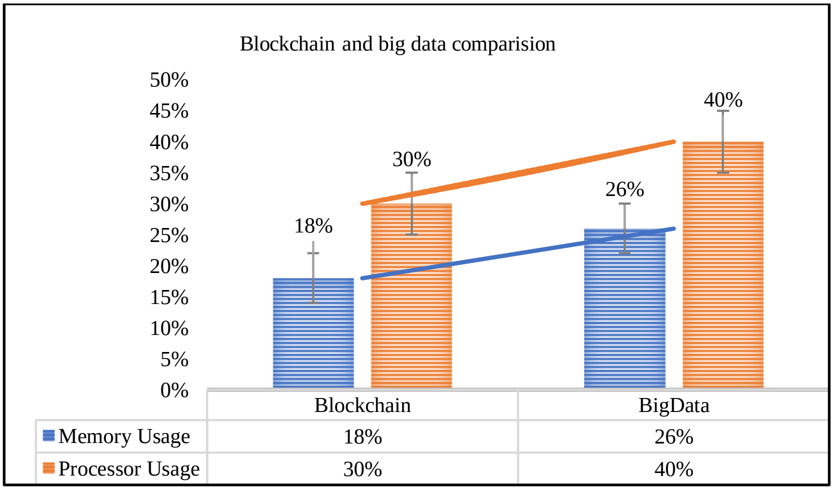


Figure 8. Memory and processor usage in blockchain and big data

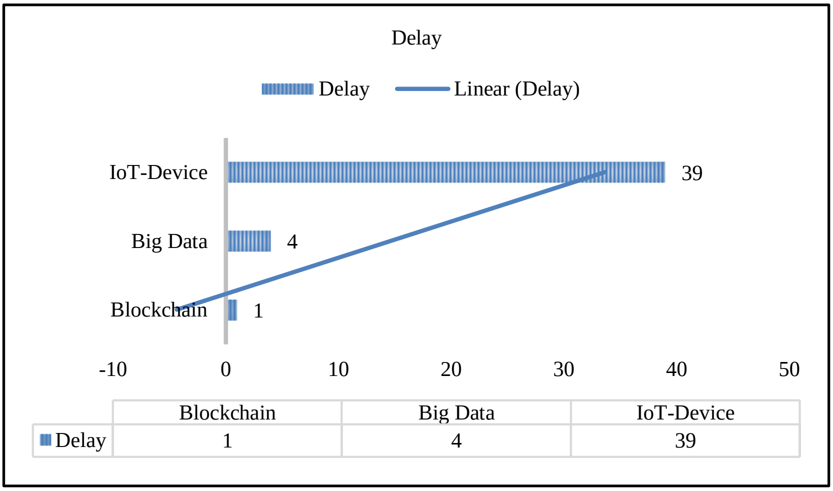


Figure 9. Transmission delay computation

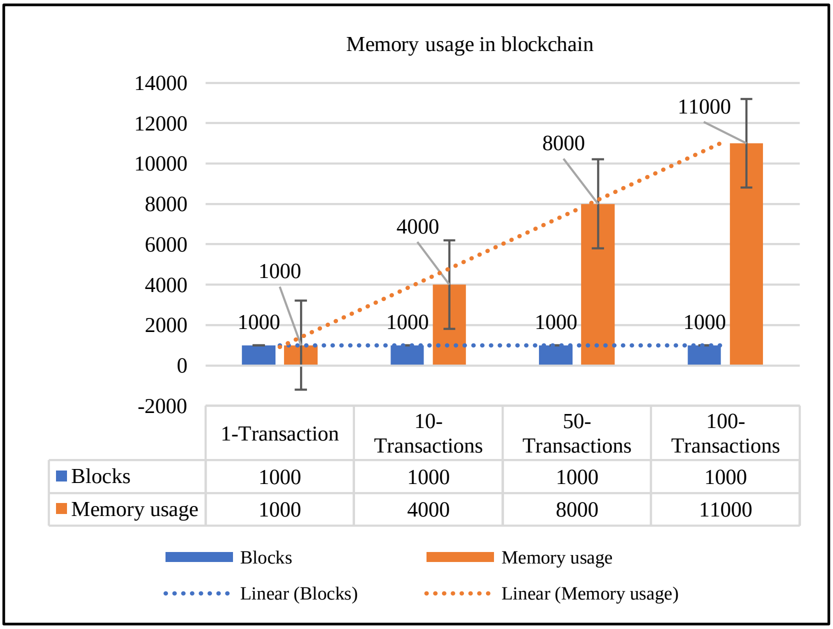
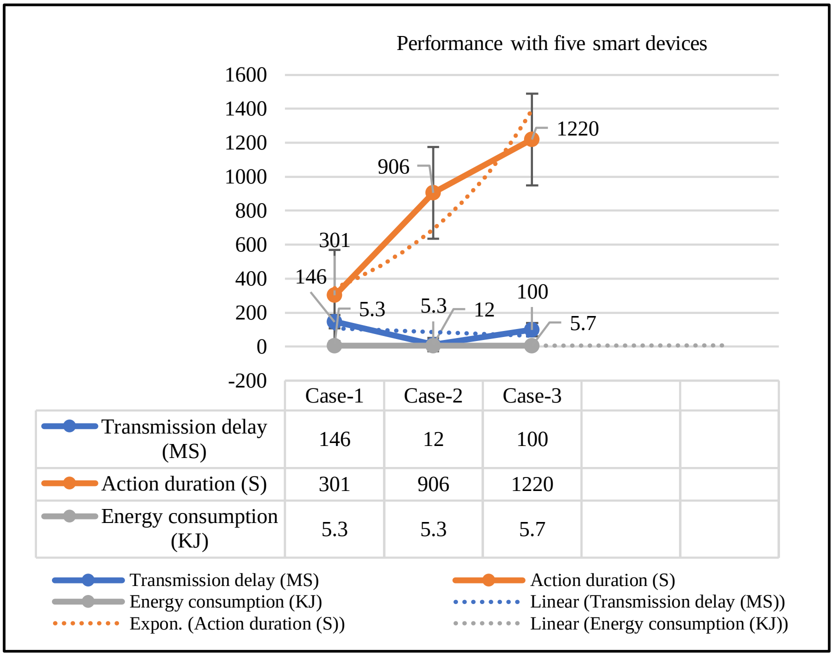
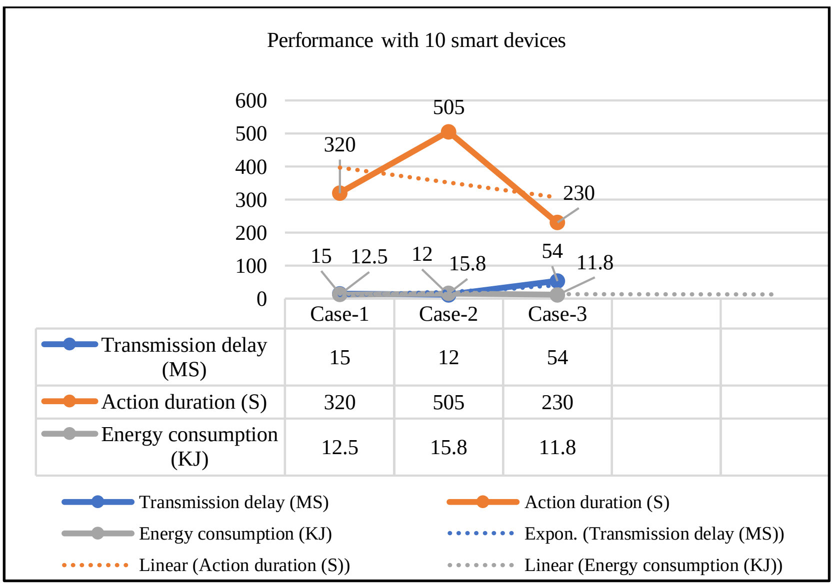


Figure 10. Memory usage and number of blocks versus transactions in blockchain



	Case-1	Case-2	Case-3		
Transmission delay (MS)	146	12	100		
Action duration (S)	301	906	1220		
Energy consumption (KJ)	5.3	5.3	5.7		

Figure 11. Performance evaluation with five smart devices in the network



	Case-1	Case-2	Case-3		
Transmission delay (MS)	15	12	54		
Action duration (S)	320	505	230		
Energy consumption (KJ)	12.5	15.8	11.8		

Figure 12. Performance evaluation with 10 smart devices in the network

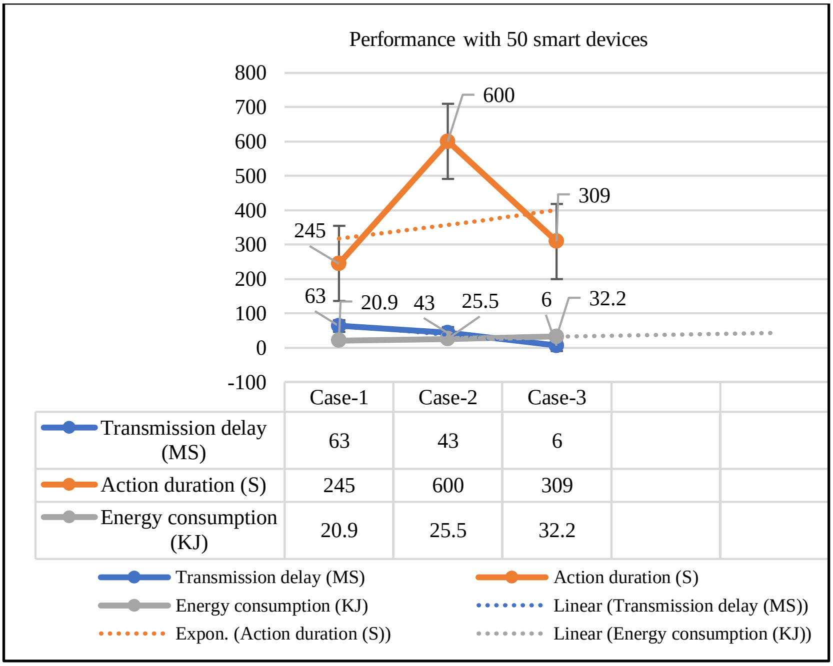


Figure 13. Performance evaluation with 50 smart devices in the network

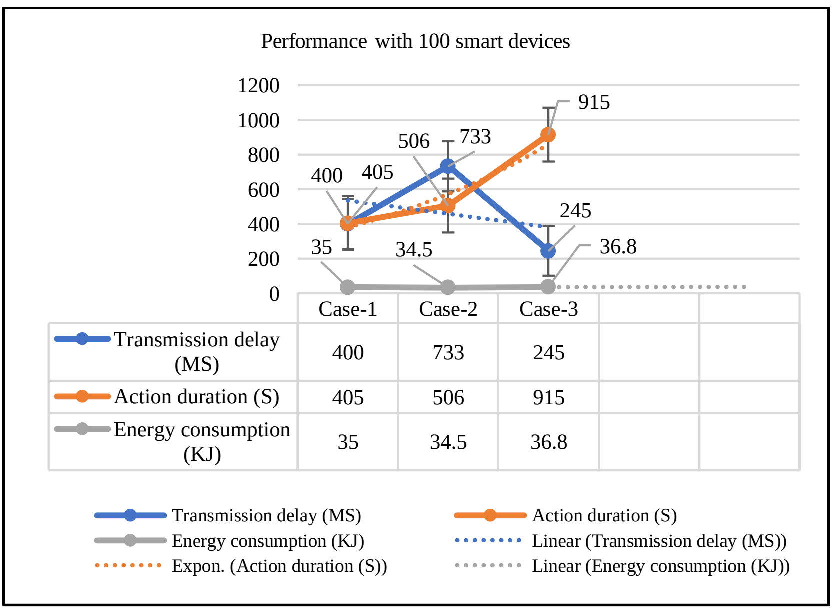


Figure 14. Performance evaluation with 100 smart devices in the network

4. Conclusion

Many researchers are trying to integrate smart city solutions with blockchain and big data to help achieve sustainable development, greater resilience, good leadership, adequate standard of living, and smart city development capabilities. Smart cities, blockchain, and big data are three new and essential approaches. Throughout this research, all terms and a variety of contexts were discussed, and some statistical inferences were defined for each. Every term has a variety of features that describe it specifically, given differing meanings. Depending on these specific features, several major advantages of the data analysis architecture and development for smart city implementations are established. In the future, researchers can enhance this research and can use it in the framework of futuristic smart cities.

5. Acknowledgement

This research was supported by the Deanship of Scientific Research, Islamic University of Madinah, Madinah, Saudi Arabia.

6. References

Alam, T. (2017). Fuzzy control based mobility framework for evaluating mobility models in MANET of smart devices. ARPN Journal of Engineering and Applied Sciences, 12(15), 4526-4538. https://doi.org/10.22541/au.159164759.95809729

Alam, T. (2020). IoT-fog-blockchain framework: Opportunities and challenges. International Journal of Fog Computing, 3(2), 1-20. https://doi.org/10.4018/IJ FC. 2020070101

Alam, T. (2019a). IoT-Fog: A communication framework using blockchain in the Internet of Things. International Journal of Recent Technology and Engineering (IJRTE), 7(6). https://doi.org/10.22541/au.159164762.24316077

Alam, T. (2019b). Middleware implementation in cloud-MANET mobility model for internet of smart devices. International Journal of Computer Science and Network Security, 17(5), 86-94. https://doi.org/10.2139/ssrn.3638980

Alam, T., & Aljohani, M. (2015). Design and implementation of an ad hoc network among Android smart devices. Proceedings of 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), Noida, India, 1322-1327.

Alam, T., & Aljohani, M. (2015). An approach to secure communication in mobile ad-hoc networks of Android devices. Proceedings of 2015 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Okinawa, Japan, 371-375.

Alam, T., & Benaida, M. (2018). CICS: Cloud-internet communication security framework for the internet of smart devices. International Journal of Interactive Mobile Technologies, 12(6), 74-84. https://doi.org/10.3991/ijim.v12i6.6776

Alam, T., & Benaida, M. (2019). The role of cloud-MANET framework in the internet of things (IoT). International Journal of Online Engineering, 14(12), 97-111. https://doi.org/10.3991/ijoe.v14i12.8338

Alam, T., Baha, R., Arshad, A., & Shamimul, Q. (2020). Distributed intelligence at the edge on IoT networks. Annals of Emerging Technologies in Computing, 4(5), 1-18. https://doi.org/10.33166/AETiC.2020.05.001

Alfa, A.S., Maharaj, B.T., Ghazaleh, H.A., & Awoyemi, B. (2018). The role of 5G and IoT in smart cities. In M. Maheswaran & E. Badidi (Eds.), Handbook of smart cities (pp. 31-54). Switzerland: Springer, Cham.

Bruyn, A.S. (2017). Blockchain an introduction. Retrieved from https://beta.vu.nl/nl/Images/werkstuk-bruyn_tcm235-862258.pdf

Chen, J., Lv, Z., & Song, H. (2019). Design of personnel big data management system based on blockchain. Future Generation Computer Systems, 101, 1122-1129. https://doi.org/10.1016/j.future.2019.07.037

Fedak, V. (2018, February 21). Blockchain and big data: The match made in heavens. Towards data science. Retrieved from https://towardsdatascience.com/blockchain-and-big-data-the-match-made-in-heavens-337887a0ce73

Google Trends (n.d). Big Data, Blockchain, Smart City (2016-2021). Retrieved from https://trends.google.com/trends/explore?date=2016-01-11%202021-01-11&q=smart%20city,blockchain,big%20data,Internet%20of%20Things

Granville, V. (2013). A synthetic variance designed for Hadoop and big data. Hadoop360. Retrieved from https://www.hadoop360.datasciencecentral.com/blog/a-synthetic-variance-designed-for-hadoop-and-big-data

Hassani, H., Huang, X., & Silva, E.S. (2019). Fusing big data, blockchain, and cryptocurrency. Switzerland: Palgrave Pivot, Cham.

Honarvar, A.R., & Sami, A. (2019). Multi-source dataset for urban computing in a smart city. Data in Brief, 22, 222-226. https://doi.org/10.1016/j.dib.2018.09.113

Karafiloski, E., & Mishev, A. (2017, July). Blockchain solutions for big data challenges: A literature review. Proceedings of IEEE EUROCON 2017-17th International Conference on Smart Technologies, Ohrid, Macedonia, 763-768.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1-14. https://doi.org/10.1007/s10708-013-9516-8

Kulkarni, P., & Akhilesh, K.B. (2020). Big data analytics as an enabler in smart governance for the future smart cities. In K.B.A. Dietmar & P.F. Möller (Eds.), Smart technologies (pp. 53-65). Singapore: Springer, Singapore.

Lueth, K.L. (2018). State of the IoT 2018: Number of IoT devices now at 7B – Market accelerating. Retrieved from https://iot-analytics.com/state-of-the-iot-update-q1-q2-2018-number-of-iot-devices-now-7b/

Bishr, A. B. (2019). Dubai: A city powered by blockchain. Innovations: Technology, Governance, Globalization, 12(3-4), 4-8. https://doi.org/10.1162/inov_a_00271

Strohbach, M., Ziekow, H., Gazis, V., & Akiva, N. (2015). Towards a big data analytics framework for IoT and smart city applications. In F. Xhafa, L. Barolli, A Barolli & P. Papajorgji (Eds.), Modeling and processing for next-generation big-data technologies (pp. 257-282). Switzerland: Springer, Cham.

Sun, Y., Song, H., Jara, A.J., & Bie, R. (2016). Internet of things and big data analytics for smart and connected communities. IEEE Access, 4, 766-773. https://doi.org/10.1109/ ACCESS.2016.2529723

Shirer, M., & MacGillivray, C. (2019). The growth in connected IoT devices is expected to generate 79.4 zb of data in 2025 according to a new IDC forecast. Retrieved from https://bwnews.pr/2M1IGNj

Tsai, C.W., Lai, C.F., Chao, H.C., & Vasilakos, A.V. (2015). Big data analytics: A survey. Journal of Big Data, 2(1), 1-32. https://doi.org/10.1186/s40537-015-0030-3