Data Allocation Mechanism for Internet-of-Things Systems With Blockchain

Wendy Yánez, Redowan Mahmud, Rami Bahsoon, *Member, IEEE*, Yuqun Zhang, *Member, IEEE*, and Rajkumar Buyya

***Abstract*—The use of Internet of Things (IoT) has introduced genuine concerns regarding data security and its privacy when data are in collection, exchange, and use. Meanwhile, blockchain offers a distributed and encrypted ledger designed to allow the creation of immutable and tamper-proof records of data at different locations. While blockchain may enhance IoT with innate security, data integrity, and autonomous governance, IoT data management and its allocation in blockchain still remain an architectural concern. In this article, we propose a novel context-aware mechanism for on-chain data allocation in IoT-blockchain systems. Specifically, we design a data controller based on fuzzy logic to calculate the Rating of Allocation (RoA) value of each data request considering multiple context parameters, i.e., data, network, and quality and decide its on-chain allocation. Furthermore, we illustrate how the design and realization of the mechanism lead to refinements of two commonly used IoT-blockchain architectural styles (i.e., blockchain-based cloud and fog). To demonstrate the effectiveness of our approach, we instantiate the data allocation mechanism in the blockchain-based cloud and fog architectures and evaluate their performance using FogBus. We also compare the efficacy of our approach to the existing decision-making mechanisms through the deployment of a real-world healthcare application. The experimental results suggest that the realization of the data allocation mechanism improves network usage, latency, and blockchain storage and reduces energy consumption.**

***Index Terms*—Blockchain, data management, fuzzy logic, Internet of Things (IoT), software architecture styles.**

Manuscript received August 17, 2019; revised December 3, 2019 and January 8, 2020; accepted February 1, 2020. Date of publication February 10, 2020; date of current version April 14, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61902169, in part by the Shenzhen Peacock Plan under Grant KQTD2016112514355531, and in part by the Science and Technology Innovation Committee Foundation of Shenzhen under Grant JCYJ20170817110848086. *(Corresponding author: Yuqun Zhang.)*

Wendy Yánez is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China, also with the School of Computer Science, University of Birmingham, Birmingham B15 2TT, U.K., and also with the Facultad de Ingeniería en Electricidad y Computación, Escuela Superior Politécnica del Litoral, ESPOL, Polythecnic University (Gustavo Galindo), Guayaquil, Ecuador (e-mail: wpy443@bham.ac.uk).

Redowan Mahmud and Rajkumar Buyya are with the Cloud Computing and Distributed Systems Laboratory, School of Computing and Information Systems, University of Melbourne, Melbourne, VIC 3053, Australia (e-mail: mahmudm@student.unimelb.edu.au; rbuyya@unimelb.edu.au).

Rami Bahsoon is with the School of Computer Science, University of Birmingham, Birmingham B15 2TT, U.K. (e-mail: r.bahsoon@cs.bham.ac.uk).

Yuqun Zhang is with the Shenzhen Key Laboratory of Computational Intelligence, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: zhangyq@sustech.edu.cn).

Digital Object Identifier 10.1109/JIOT.2020.2972776

I. Introduction

THE RAPID evolution and adoption of the Internet of Things (IoT) have contributed to advancements in our society, leveraging the potential of data and smart environments [1], [2]. According to Gartner [3], 26 billion IoT devices will be connected to the Internet by 2020, while CISCO [4] predicts that by the same year, the number of interconnected devices will reach around 50 billions. With the growing number of devices and a huge volume of data gathered by them, data management and its security become key concerns in IoT systems and services [5]–[9]. In general, IoT is defined as a network of devices that collect data from the environment and communicate with each other to enable advanced applications across different domains, e.g., healthcare, manufacturing, environmental monitoring, etc. [10]. The traditional IoT systems rely on the cloud for data processing and storage, but it could become a single point of failure as the number of devices increases in the network as well as lead to security attacks on time-sensitive IoT data.

Recently, blockchain has emerged as a promising technology that provides a distributed ledger where transactions are protected by cryptography and verified in a Peer-to-Peer (P2P) network to enhance decentralization and secure data sharing [11]. Some influential companies, e.g., IBM, emphasize the role of blockchain in realizing the potential of IoT systems and the democratization of its network of things [12]. Blockchain offers a distributed storage where data gathered by IoT devices can be recorded in an immutable and verifiable manner without the need of a third party [13]. With such features, it is possible to track all the actions executed on IoT networks over time in order to trigger timely decisions for the purpose of regulatory compliance and system operation. Overall, the integration of blockchain in fog and cloud infrastructures is expected to offer dependable and trustworthy hosting environments for IoT data transactions and secure data sharing [14], [15].

Despite the increasing interest in blockchain and IoT, there is still a lack of systematic approaches that consider IoT data management and its allocation as one of the significant architectural design decisions for developing IoT-blockchain systems [5]. Bridging this gap is essential for advancing in the adoption of blockchain in IoT and for facilitating the design and deployment of IoT-blockchain architectural styles. In this article, we draw on a healthcare case study provided in [16] as an example to identify the significant architectural requirements of a data-centric IoT-blockchain architecture. To

address these requirements, we propose a novel context-aware mechanism for supporting on-chain data allocation based on multiple context parameters, e.g., data, network, and quality. Data refers to raw data gathered by IoT devices, the network corresponds to the number of sharing points interested in the IoT data, and quality relates to accurate measurement of the device themselves.

The proposed mechanism mainly relies on a data controller based on fuzzy logic to support data allocation decisions out of the imprecision and ambiguity of multiple input parameters. Specifically, the data controller translates crisp sensor data (i.e., bits) to fuzzy inputs (i.e., severe, moderate, etc.) using the domain expert membership functions to get the fuzzy output. This fuzzy output is mapped to a machine-readable output using the defined membership functions and serves as the threshold value called the Rating of Allocation (RoA) to decide on-chain allocation. To demonstrate the flexibility of our approach, we instantiate the mechanism in two commonly used IoT-blockchain architectural styles for integrating blockchain and IoT, i.e., blockchain-based cloud and fog [15], [17]. Leveraging blockchain in the fog and the cloud, it provides extra security to the two computing environments by ensuring data immutability, traceability, and integrity [16]. However, the design and realization of the data allocation mechanism lead to refinements of the existing architectural styles which should consider the QoS requirements of IoT systems and the constraints imposed by the hosting environments, e.g., fog and cloud. To this end, we envision a four-tier abstraction, i.e., the IoT tier, the data controller tier, the fog tier, and the cloud tier where the data controller tier is introduced between the IoT tier and the fog tier. The data controller tier enables the data allocation mechanism to decide which data need to be stored within the blockchain embedded in the fog or the cloud or allocated off-chain (e.g., cloud database).

We compare our approach against some existing decision-making mechanisms, such as logistic regression [18], [19] and decision tree [20], [21] to evaluate how they perform in terms of CPU usage and execution time. We also evaluate the refined blockchain-based cloud and fog architectures in the above healthcare case study by using the FogBus framework [16]. First, the mechanism is run on the refined blockchain-based cloud and then executed on the refined blockchain-based fog. The evaluation mainly focuses on enabling and disabling the data allocation mechanism in the two IoT-blockchain architectural styles and compares their performance in terms of blockchain size, latency, energy consumption, and network usage. The results show that latency is reduced by 36% in the refined blockchain-based cloud and about 27% in the refined blockchain-based fog. Similarly, energy consumption is reduced by averagely 28% in the refined blockchain-based cloud and fog. Moreover, the network usage is reduced by 32% in the refined blockchain-based fog and 24% in the refined blockchain-based cloud.

The main contributions of this article are summarized as follows.

1) We develop a novel context-aware data allocation mechanism to determine the RoA value based on multiple context parameters and support on-chain data allocation.
2) We illustrate how the design and realization of the data allocation mechanism lead to refinements in the two state-of-the-art IoT-blockchain architectural styles, i.e., blockchain-based cloud and fog.
3) We evaluate the refined blockchain-based cloud and fog by applying them to a healthcare case study using FogBus. The experimental results suggest significant improvements in data transaction latency, network usage, energy consumption, and blockchain storage usage.

The remainder of this article is organized as follows. Section II introduces the background of IoT and blockchain. Section III presents a motivation scenario and its architectural requirements for developing IoT-blockchain systems. Section IV models the proposed data allocation mechanism. Section V introduces the refined blockchain-based cloud and fog architectures. Section VI provides an illustrative example that supports the implementation of the data allocation mechanism in the two commonly used IoT-blockchain architectural styles. Section VII conducts a set of experiments to evaluate the effectiveness of the data allocation mechanism and compares our approach to other existing decision-making mechanisms. Section VIII summarizes the related work in fuzzy logic and IoT. Section IX presents the envisioned challenges and possible future research and Section X concludes this article.

II. Background

In this section, we briefly introduce IoT and blockchain.

A. Internet of Things

The IoT consists of a network of devices that monitor, collect, and exchange data over the Internet to enable intelligent applications, e.g., healthcare, manufacturing, smart cities, transportation, etc. [22]–[24]. IoT devices range from sensors and actuators with limited CPU, memory, and battery resources to devices with advanced computing capabilities, e.g., mobile phones, vehicles, home appliances, etc. [6], [10]. Current IoT systems rely on centralized cloud servers for data processing and storage, but the transfer of a huge volume of data to the cloud could lead to high latency and bandwidth consumption across the network. In addition, data collected by IoT devices include sensitive and critical information that could be manipulated, altered, and tampered by untrusted parties in cloud infrastructures [25]. With the advances in IoT development, fog extends cloud functions at the edge of IoT networks to ensure quick processing and short-term storage on time-sensitive IoT data [26], [27]. Specifically, fog enables a distributed and heterogeneous network of devices called fog nodes (e.g., controllers, switches, gateways, and embedded servers) that process and analyze IoT data close to the devices instead of sending it to the cloud [28]. This minimizes latency and bandwidth consumption and keeps sensitive data inside the network [29], [30]. Thus, the use of fog and/or cloud depends on the specific requirements, constraints, and tradeoffs imposed by IoT systems [31].

B. Blockchain

Blockchain is emerging as a distributed storage service supported by a P2P network where transactions are protected with cryptography and validated in consensus [32], [33]. Leveraging on cryptography, blockchain ensures data integrity, authenticity, traceability, and accountability in data exchange between untrusted devices in the network [13]. The decentralized and distributed nature of blockchain can play a significant role in how IoT devices communicate directly with eachother or with minimal human intervention. With blockchain in IoT, device data can be recorded as immutable and tamper-proof transactions over time and shared among devices in the network [11], [32].

Basically, blockchain consists of a list of blocks where each block stores a set of transactions. Each time a transaction is generated, it is signed using public-/private-key pair and broadcasted to all the nodes in the blockchain network. Nodes receiving the transaction verify the signature attached to it and validate each transaction in consensus by following a mining process [14], [32]. During mining, nodes known as miners allocate verified transactions into cryptographically secure blocks by solving a consensus problem. Upon the receipt of a new block, the miners append the new block with verified transactions at the end of the chain. Each new block contains a hash of the previous block to enhance the consistency of the ledger.

III. Motivation Example and Requirements

This section introduces a case study as the motivation example to identify the architectural significant requirements to consider when integrating blockchain and IoT and supports on-chain data allocation.

A. Sleep Apnea Case Study

We use a sleep apnea case study provided in [16] to motivate our approach and develop a data allocation strategy for IoT data management in the blockchain. We choose this example in the healthcare domain due to its applicability and similar features to other IoT cases in terms of scale and size [34].

Assume Andrew, an 85-year-old male, is experiencing snoring and excessive daytime sleepiness which could be diagnosed as a sleep apnea disorder. Sleep apnea occurs when normal breathing is interrupted or stopped completely during sleep, resulting in high-blood pressure, heart disease, brain stroke, etc. [35]. Andrew's doctor recommends using a pulse oximeter (e.g., wearable device) on his finger to monitor the heartbeat and oxygen saturation in his blood and determine the sleep apnea level. According to the number of apnea events recorded per hour of the sleep study, the apnea-hypopnea index (AHI) is classified as follows.

1) *Mild:* AHI $\geq$ 5, but $<$ 15 per hour.
2) *Moderate:* AHI $\geq$ 15, but $<$ 30 per hour.
3) *Severe:* AHI $\geq$ 30 per hour.

Due to the limited capabilities in the oximeter, it is connected to Andrew's smartphone to forward the collected data to the fog and cloud infrastructures. We assume architects implement a blockchain in both environments (e.g., fog and/or cloud) to protect Andrew's health data and secure share it with healthcare providers (i.e., doctors, hospitals, pharmacies, laboratories, health insurance companies, etc.). If Andrew's is moved from one hospital to another, the uncertainty in the normal range for test results could make difficult for medical staff to diagnose the disease. In addition, the incomplete and missing information on his health history makes decisions more complex and uncertain.

B. Architectural Significant Requirements

We use the above example to identify a number of architectural significant requirements that support the development of a data-centric approach for supporting on-chain data allocation. In particular, we present the requirements for the healthcare scenario regarding uncertainty, imprecision, vagueness, fuzziness, data incompleteness, and nonbinary representation of such issues. We argued that IoT systems can be subject to a variety of uncertainties in their operation environment, such as changes in traffic network and interference [36], [37]. These uncertainties could result in incomplete, imprecise, and missing information that makes it difficult to offer accurate decision support [38], [39]. On the other hand, many real-world problems essentially demand multifactor consideration at the same time before making decisions [40]. In particular, there can be a number of real-world scenarios that cannot be simply analyzed/depicted by a set of binary values if they depend on various factors for making decisions [41], [42]. For instance, in a cold supply chain system, instead of simple binary descriptions "cold or hot," indoor and outdoor conditions, such as various temperature/humidity levels, need to be considered in order to optimally maintain frozen food products under the complex threshold policies [43]. Similarly, in our sleep-apnea example, the oximeter data (i.e., heart rate and oxygen saturation in the patient's blood) could not be sufficient to understand its criticality or sensitivity levels to maintain its management and allocation soundness. In addition, we rely on blockchain to enhance the security and privacy of health data and securely share it among interested healthcare providers [11], [32]. We summarize the requirements that support adopting fuzzy logic and blockchain via the healthcare example in our approach as follows.

1) *Requirement 1 (R1):* This approach copes with uncertainty and imprecise information. Patients and/or his/her family often fail to precisely express their symptoms and rather use ambiguous terms that could lead to many suboptimal and even incorrect medical decisions.
2) *Requirement 2 (R2):* This approach copes with missing and incomplete information from sensors, caused by the heterogeneous hardware and software in IoT devices and dynamic traffic in IoT networks (i.e., devices join and leave the network).
3) *Requirement 3 (R3):* This approach considers the cases with more than one decision-making dimensions. In particular, multiple sensor readings as well as environmental information need to be collected to diagnose patient conditions and deliver accurate treatment.

4) *Requirement 4 (R4):* This approach relies on decentralized infrastructures for secure data storage and data sharing among interested parties. In particular, IoT systems tend to shift from centralized infrastructures to record data in a decentralized fashion and empower users with control over their records.

The application of fuzzy logic and blockchain fits the above requirements on the data-value uncertainty *R1, R2, R3*, and the security requirements of critical IoT applications *R4*. To meet these requirements, we design a data controller based on fuzzy logic that extracts multiple context parameters of each data request, i.e., data, network, and quality to calculate the RoA. This value gives us insights about the sensitivity of IoT data in order to decide which data request needs to be stored in the blockchain or allocated off-chain (e.g., cloud database). In the context of our study, the context parameters, i.e., data, network, and quality are used as inputs of the proposed mechanism where *data context* refers to the sleep apnea levels, i.e., mild, moderate, and severe [16] which could be used by malicious parties to infer users profile. *Network context* relates to the number of sharing points in the healthcare network interested in the collected data, e.g., doctors, hospitals, pharmacies, laboratories, healthcare insurances, etc. [44]. *Quality context* corresponds to device accuracy measurements that need to be protected to guarantee a trustworthy medical analysis [22]. We compute the three parameters to derive the RoA value and determine whether a particular data request needs to be allocated within the blockchain or kept off-chain. In addition, we enrich two existing IoT-blockchain architectural styles (i.e., blockchain-based cloud and fog) with context-aware data control and data management capabilities to support on-chain data allocation. Specifically, we introduce a data controller tier between the IoT and the fog tiers to handle on-chain or off-chain data allocation decisions based on the RoA value.

IV. Context-Aware Data Allocation Mechanism Based on Fuzzy Logic

In this section, we provide a brief introduction to Fuzzy Logic and motivate our approach. Next, we explain the proposed context-aware data allocation mechanism and the calculation of the RoA value for supporting data allocation.

A. Fuzzy Logic

Fuzzy logic is an artificial intelligence (AI) technique that uses linguistic variables to imitate human thinking and enable decision making in real-time systems [45]. This approach aims solving problems that are difficult to formulate mathematically due to imprecise or non-numerical information, such as "very cold" or "not very satisfied" [46]. Unlike classical one-to-one input-to-output control strategy, fuzzy logic makes decisions out of many-to-one and many-to-many input-to-output control [47] by using fuzzy sets and rules [48], [49]. The fuzzy-logic process consists of the three stages: 1) fuzzification; 2) inference rules; and 3) defuzzification as shown in Fig. 1.

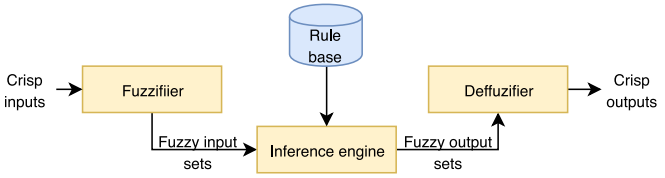


Fig. 1. Fuzzy-logic process [45].

1) *Fuzzification* converts crisp input data gathered by sensors (i.e., bits) to a fuzzy input set of linguistic terms using the membership functions.
2) *Inference* applies a set of IF-THEN rules defined by domain experts to derive the fuzzy output.
3) *Defuzzification* maps the fuzzy output to a machine-readable crisp output by using the defined membership functions.

B. Rationale Behind the Adoption of Fuzzy Logic and Blockchain

IoT networks are subject to changes in the operation contexts, such as dynamic traffic and interference [36], [37], which can lead to uncertainties regarding data values, management, and allocation. These uncertainties mainly caused by the network volatility and connectivity changes can result in several issues, e.g., data inconsistency, incompleteness, imprecision, and/or vagueness [38], [39]. If we only apply coarse-grained representations, i.e., true and false, for depicting system features/outcomes, we would end up with superficial understandings/decisions [41] of the systems. In particular, the binary logic deals with two possible values, 0 (false) and 1 (true). For instance, to make an air-conditioner decision based on the indoor temperature, if the temperature hits above 30 °C, then turn on the cooler mode. Otherwise, if the temperature hits below 18 °C, then turn on the heater mode. In general, we can infer that the binary logic is suitable for the scenarios where the solutions are made binary under the policies that exhibit certainties with reliable sensing and affirmative values, such that data management and its outcome can be reliably predicted [41]. On the other hand, in dynamic/adaptive systems, e.g., AI-based systems, the allocation policies and decisions could be more complex to set prior to execution and thus infeasible to simply apply "true or false" to depict system states [40]. Additionally, system states can be varying under different contexts, e.g., what is true under one context may be false under another [39], [42]. To address these uncertainties and make more granular decisions, we apply fuzzy logic that interprets and reasons about multiple states at a time compared with the traditional logic that only deals with two states [40], [41]. Specifically, we propose a context-aware mechanism based on fuzzy logic that considers context information to optimize on-chain allocation decisions, given changes in operation context and internal dynamics in IoT systems supported with blockchain. Our fuzzy strategy considers multiple context parameters, such as data, network, and quality to decide on which data need to be recorded on-chain or kept in external storage (i.e., cloud database).

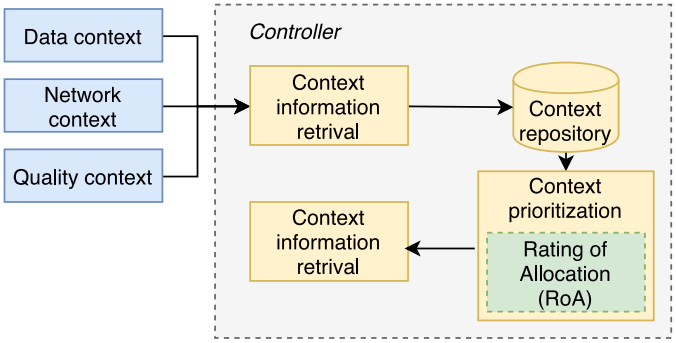


Fig. 2. Data controller components.

Moreover, sensor data are likely to include sensitive and critical information that could be manipulated and altered by untrustworthy service providers in the cloud and lead to data loss and financial damage [28]. When blockchaining data sensed by IoT devices, architects should also consider that the computation and storage space in blockchains still remain limited [32]. For instance, public blockchains can handle on average 3–20 transactions per second while VISA[1] can support around 1700 transactions per second. Therefore, it is essential to develop an efficient data allocation mechanism that copes with uncertainty in data management for IoT systems supported with blockchain. We propose a context-aware mechanism based on fuzzy logic that extracts context information from data, network, and quality to make optimal on on-chain allocation decisions. Leveraging on fuzzy logic, we aim at minimizing the risk of uncertainty, vagueness, and interpretation of incomplete data and offering appropriate allocation decisions in IoT systems supported with blockchain.

C. Envisaged Contexts

Context is defined as the computational representation of any information that can characterize the status of an entity, e.g., user, device, software application, or any other object that handles the interaction between users and services [50]. In this article, the context parameters, i.e., data, network, and quality are modeled as follows.

1) *Data context* is represented as α and refers to the device data, e.g., heart beat rate, oxygen saturation, etc. [16].
2) *Network context* is represented as β and corresponds to the number of the sharing points interested in the collected data [44].
3) *Quality context* is represented as ω and refers to the device accuracy measurements [51].

D. Data Controller Structure

Fig. 2 shows the components of the data controller, such as context information retrieval, context repository, context prioritization, and context allocation decision.

1) *Context information retrieval* extracts the context parameters, i.e., data, network, and quality from the IoT devices and the sharing points interested in the collected data.

[1]https://usa.visa.com/run-your-business/small-business-tools/retail.html

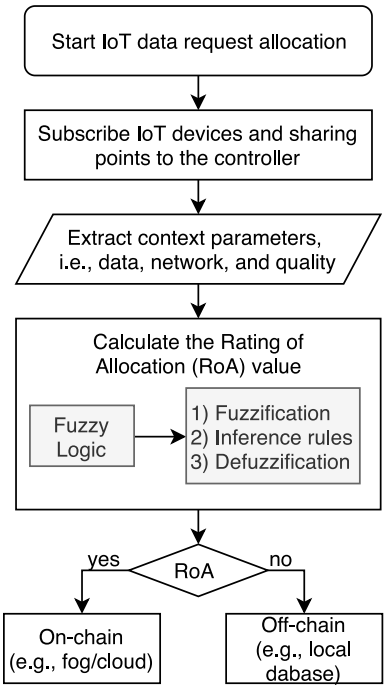


Fig. 3. Flowchart of the data allocation mechanism.

TABLE I
NOTATIONS

Symbol	Definition
R	Set of all data requests d_r.
E_{d_r}	Context parameters within a data request d_r.
α	Data context.
β	Network context.
ω	Quality context.
$U_i^{d_r}$	Context parameters (value) of i_{th} for application $d_r \in R$.
δ_{d_r}	Rating of Allocation (RoA) of a data request d_r.
μ_i	Fuzzy membership function for any context parameters E_{d_r}.
F_c	Fuzzy output set for RoA calculation.
f_{d_r}	Fuzzy output in fuzzy output set F_c.
ϕ^{d_r}	Singleton value for a fuzzy output $f_{d_r} in F_c$.
μ_o	Membership function for any fuzzy output f_{d_r} in RoA calculation.
D_s,S_p,D_q	Fuzzy sets for data sensitivity, sharing points, and data quality.

2) *Context repository* temporarily stores the retrieved context parameters before moving them to the context prioritization component.
3) *Context prioritization* computes the RoA value based on the context parameters and sharing points.
4) *Context allocation decision* uses the RoA value as a threshold measurement to determine which data request needs to be allocated within the blockchain or stored off-chain.

Fig. 3 illustrates the proposed data allocation mechanism and Table I defines the relevant notations.

The data allocation mechanism is initialized by subscribing IoT devices and sharing points interested in the collected data to the data controller. After the subscription, the *context information retrieval* extracts context parameters $E_{d_r} \in \{U_\alpha^{d_r}, U_\beta^{d_r}, U_\omega^{d_r}\}$ of each data request d_r in a set of data requests R. The context parameters in E_{d_r} are stored in the *context repository* and forwarded to the *context prioritization* component for processing and analysis. Since each context parameter in E_{d_r} uses different ranges and scales, the following

TABLE II
SCOPE OF CONTEXT PARAMETERS

Parameters	Values
$[\gamma_\alpha, \lambda_\alpha]$	5 to 40
$[\gamma_\beta, \lambda_\beta]$	1 to 5
$[\gamma_\omega, \lambda_\omega]$	0.1 to 1

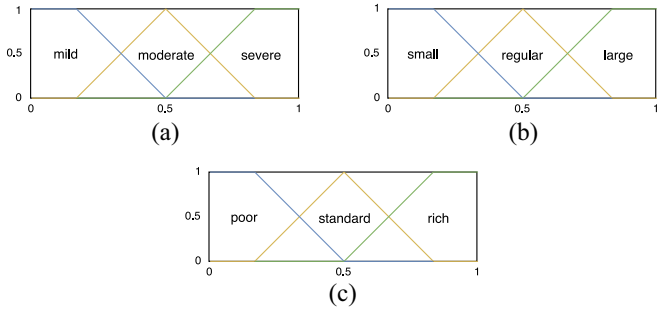


Fig. 4. Membership functions of the context parameters. (a) Data sensitivity. (b) Sharing points. (c) Data quality.

equation is used to ensure that numerical values of context parameters are normalized in the range 0–1:

$$\overline{U_i^{d_r}} = \frac{U_i^{d_r} - \gamma_i}{\lambda_i - \gamma_i}. \tag{1}$$

$U_i^{d_r}$ corresponds to the ith numerical value of a data request d_r defined in the range $[\gamma_i, \lambda_i]$ which is set according to the range of context parameters defined in Table II. For instance, the data context (α) derived from the sleep apnea level is represented in a range 5–40 [16]; network context (β) related to the number of the sharing points interested in the collected data are represented in a range 1–5 [44]; and quality context (ω) referred to the device accuracy is represented in a range 0.1–1 [51]. If numerical values of any context parameters do not fit within the defined ranges, the data request is discarded from placing in the blockchain and allocated in external storage.

Next, a fuzzy-logic approach is used to build a data controller that calculates the RoA value of each data request represented as δ_{d_r} based on the normalized context parameters in E_{d_r}. In the fuzzification phase, the normalized value $\overline{U_i^{d_r}}$ of any context parameter in E_{d_r} is converted into a fuzzy input set using the corresponding membership functions μ_i. Here, the membership functions of the collected context parameters, e.g., data, network, and quality are applied to three fuzzy sets, i.e., data sensitivity, sharing points, and data quality as illustrated in Fig. 4. Each fuzzy set is defined within a normalized range 0–1 as follows.

1) *Data Sensitivity:* $D_s \in$ {Mild, Moderate, Severe}.
2) *Sharing Points:* $S_p \in$ {Small, Regular, Large}.
3) *Data Quality:* $D_q \in$ {Poor, Standard, Rich}.

The membership degree $\mu_i(\overline{U_i^{d_r}})$ for any normalized value $U_i^{d_r}$ on the corresponding fuzzy set can be graphically represented as triangular waveform, trapezoidal waveform, etc. [45]. Here, the trapezoidal waveform is used to represent the dynamic variation of the context parameters in the IoT system. Each membership function has a grade from 0 to 1 at each end point and uses a label to identify its condition.

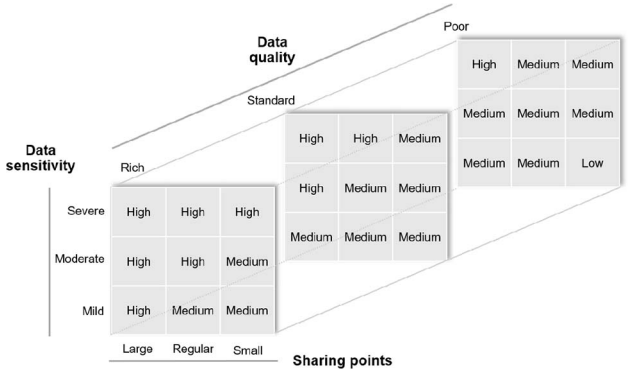


Fig. 5. Fuzzy rules for RoA calculation.

In the fuzzy inference phase, the data controller evaluates the fuzzy input data according to the domain expert IF-THEN rules, where IF captures the system's knowledge using a condition and THEN derives the corresponding fuzzy output as a conclusion. These domain-specific rules allow comparing the relation between multiple input and output parameters. Fig. 5 illustrates a set of fuzzy rules with their corresponding fuzzy output set for calculating the RoA value defined as $F_c \in$ {Low, Medium, High}.

The following are some representative examples of the fuzzy rules to determine the fuzzy output $f_{d_r} \in F_c$ for a data request d_r used by the data controller to calculate the RoA value.

1) **IF** data sensitivity (α) is severe AND sharing points (β) is large AND data quality (ω) is rich, **THEN** RoA is high.
2) **IF** data sensitivity (α) is normal AND sharing points (β) is regular AND data quality (ω) is rich, **THEN** RoA is medium.
3) **IF** data sensitivity (α) is mild AND sharing points (β) is small AND data quality (ω) is poor, **THEN** RoA is low.

In addition to evaluating the rules, the inference phase also combines the results of each rule to determine the fuzzy output. According to the fuzzy rules, severe data sensitivity (e.g., rigid parameter) is given higher weight compared to regular sharing points and standard data quality (e.g., relaxed parameters), since data sensitivity could be used to infer users profile and perform malicious attacks. As a result, the RoA value becomes more aligned to the data sensitivity parameters than the other relaxed parameters. The context parameters in a data request are logically linked through the AND operator to deliver the fuzzy output. This operator represents the intersection of membership functions whose values for each context parameter is defined as the *minimum* of individual membership functions [45]. The following equation is used to calculate the membership function of fuzzy output $\mu_o(f_{d_r})$ for a data request d_r:

$$\mu_o(f_{d_r}) = \min\left(\mu_\alpha\left(\overline{E_\alpha^{d_r}}\right), \mu_\beta\left(\overline{E_\beta^{d_r}}\right), \mu_\omega\left(\overline{E_\omega^{d_r}}\right)\right). \tag{2}$$

Based on the context input parameters, multiple rules can be triggered at the same time which requires combining the membership functions of the associate fuzzy output to derive

the final result. In the defuzzification phase, the fuzzy output is mapped to a crisp machine-readable output using the defined membership functions. Here, the RoA value δ_{d_r} of a data request d_r is calculated by combining the membership functions of the fuzzy output and using a set of singleton values to distinguish different outputs. For each fuzzy output f_{d_r}, there is a Singleton value $\phi_k^{f_{d_r}}$ that is defined as the maximum rate of the a data request for the fuzzy output f_{d_r}. The following equation calculates the deffuzified RoA value denoted as δ_{d_r} using the discrete center of gravity method:

$$\delta_{d_r} = \frac{\sum_{n=1}^{n=k} \mu_o\left(f_{d_r}^k\right) * \phi_k^{f_{d_r}}}{\sum_{n=1}^{n=k} \mu_o\left(f_{d_r}^k\right)}. \tag{3}$$

Here, δ_{d_r} corresponds to the RoA value for a data request d_r after applying fuzzy logic on the context parameters in E_{d_r}. Next, δ_{d_r} is used by the *context allocation decision* component to derive the allocation decision for a data request d_r. In particular, when the context parameters of a data request E_{d_r} are higher than δ_{d_r}, then it is allocated within the blockchain, otherwise it is stored off-chain (e.g., cloud database).

V. Integrating Data Allocation Mechanism With IoT-Blockchain Architectures

Many studies have focused on the implementation of blockchain in fog and cloud environments to enhance their security in terms of data immutability, traceability, and integrity [15], [17]. In [17], a blockchain-based cloud framework is proposed where cloud servers (i.e., application servers, data servers, etc.) become trusted nodes that support IoT data transactions in a distributed and secure manner. Furthermore, in [15], a blockchain-based fog is designed to ensure fog nodes are tamper-proof and data on them cannot be manipulated or altered by untrusted parties. Despite the interest of embedding blockchain either in fog or cloud, there is still a need for enhancing blockchain-based cloud and fog architectures with data management and allocation capabilities in order to alleviate the storage capacity of blockchain. To this end, we propose a data allocation mechanism that calculates the RoA value of each data request based on multiple context parameters and decides its allocation on-chain or off-chain. The implementation of the mechanism leads to refinements in the IoT-blockchain architectural styles that should reflect the way the mechanism is integrated into them considering the QoS requirements and the constraints of cloud and fog environments. To handle these refinements, we introduce a data controller tier between the IoT tier and the fog tier which handles data allocation decisions as shown in Fig. 6, with the details of all the tiers illustrated as follows.

1) *IoT tier* consists of sensors and actuators that perceive information from the environment and act on the collected data. As many IoT devices have limited computing capabilities to preprocess real-time data, they are connected to proximate gateways to transmit the collected data to the upper tiers.
2) *Data controller tier* acts as a broker interface between the IoT tier and the fog tier and consists of a network of gateways nodes that implements the data controller logic. The data controller extracts context parameters of each data request, e.g., data, network, and quality to calculate the RoA value and determine its allocation within the blockchain or off-chain, e.g., cloud database.
3) *Fog tier* enables a network of distributed nodes with advanced capabilities (e.g., gateways, switches, local servers, etc.) that ensure quick-processing and short-term storage close to where the data are collected which reduces the amount of data sent to the cloud.
4) *Cloud tier* enables a centralized and scalable platform with significant processing and storage resources to support the deployment of IoT applications with minimal cost.

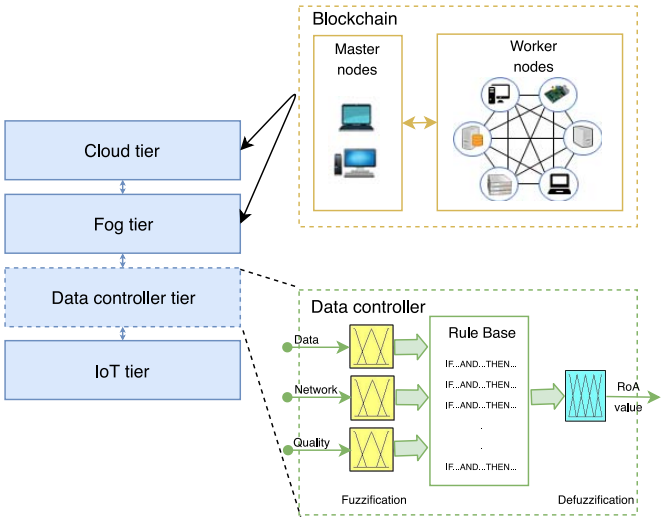


Fig. 6. Enrichment of the IoT-blockchain architectural styles.

As our approach relies on enriching two commonly used IoT-blockchain architectural styles, i.e., blockchain-based cloud and fog, the blockchain network is designed as follows. *Blockchain* consists of a network of heterogeneous nodes in terms of processing, storage, and energy resources that play different roles in the architecture, e.g., master and worker nodes. This tier can be implemented across diverse computing infrastructures, e.g., fog and cloud to reduce the overhead in the network and enable secure data sharing in both the hosting environments.

1) *Master nodes* receive data requests as transactions from the data tier and discover worker nodes that can process them in a distributed manner. Moreover, the master nodes generate a public-/private-key pair for signing the received transactions before broadcasting them in the blockchain network. These nodes also create blocks to store confirmed transactions and calculate the hash of each block to append it to its own chain.
2) *Worker nodes* use the public key provided by master nodes to verify whether the received transaction is coming from a legitimate source. Once the transaction is verified, it is validated in consensus following a mining process to be considered confirmed.

VI. Illustrative Example

We explain step by step the proposed data allocation mechanism using the healthcare example described in Section III-A.

TABLE III
PARAMETERS OF DATA REQUESTS

Id	α	β	ω	δ
Req1	$E^1_\alpha = 40$	$E^1_\beta = 5$	$E^1_\omega = 1$	8.31
	$\overline{E^1_\alpha} = 1$	$\overline{E^1_\beta} = 1$	$\overline{E^1_\omega} = 1$	
Req2	$E^4_\alpha = 30$	$E^4_\beta = 3$	$E^4_\omega = 0.8$	8.07
	$\overline{E^4_\alpha} = 0.7436$	$\overline{E^4_\beta} = 0.5$	$\overline{E^4_\omega} = 0.7778$	
Req3	$E^3_\alpha = 35$	$E^3_\beta = 3$	$E^3_\omega = 1$	8.27
	$\overline{E^3_\alpha} = 0.8718$	$\overline{E^3_\beta} = 0.5$	$\overline{E^3_\omega} = 1$	
Req4	$E^5_\alpha = 15$	$E^5_\beta = 1$	$E^5_\omega = 1$	5
	$\overline{E^5_\alpha} = 0.359$	$\overline{E^5_\beta} = 0$	$\overline{E^5_\omega} = 1$	
Req5	$E^2_\alpha = 30$	$E^2_\beta = 3$	$E^2_\omega = 0.5$	7.52
	$\overline{E^2_\alpha} = 0.7436$	$\overline{E^2_\beta} = 0.5$	$\overline{E^2_\omega} = 0.4444$	
Req6	$E^6_\alpha = 38$	$E^6_\beta = 3$	$E^6_\omega = 0.6$	8.19
	$\overline{E^6_\alpha} = 0.9487$	$\overline{E^6_\beta} = 0.5$	$\overline{E^6_\omega} = 0.5556$	
Req7	$E^7_\alpha = 31$	$E^7_\beta = 4$	$E^7_\omega = 0.7$	7.91
	$\overline{E^7_\alpha} = 0.7692$	$\overline{E^7_\beta} = 0.75$	$\overline{E^7_\omega} = 0.6667$	
Req8	$E^8_\alpha = 20$	$E^8_\beta = 5$	$E^8_\omega = 1$	7.694
	$\overline{E^8_\alpha} = 0.4872$	$\overline{E^8_\beta} = 1$	$\overline{E^5_\omega} = 0.4444$	

Here, the data controller receives between 20 and 200 data signals per data request from IoT devices and sharing points at any time *t*. For each data request, the data controller extracts context parameters, i.e., data, network, and quality where data context refers to the sleep apnea level gathered from the pulse oximeter in a range of 5–40, network context corresponds to the number of the sharing points interested in the user data (i.e., doctors, hospitals, pharmacies, insurance companies, etc.) in a range of 1–5, and quality context relates to accuracy value measurement provided by the device itself in a range of 0.1–1. Next, the data controller normalizes the context parameters in order to calculate the RoA value and support on-chain allocation decisions accordingly. Table III illustrates representative examples of context parameters serving as an input of the data controller along with their normalized values and RoA value outputs. In addition to these parameters, the singleton values are defined as $\phi_{\text{High}} = 10, \phi_{\text{Medium}} = 5$, and $\phi_{\text{Low}} = 2$ to make a clear distinction between intermediate levels of the fuzzy output set F_c defined as Low, Medium, and High. These values are associated to the degree of membership of a particular fuzzy set and defined in the same order as described in Section IV-D.

The result in Req1 reveals that RoA is high (8.31) when data sensitivity is 40, sharing points is 5, and data quality is 1. From Req5, we realize that RoA value is high (7.52) when data sensitivity is 30, sharing points is 3, and data quality is 0.5. However, in Req4 the RoA value is low (5) when data sensitivity is 15, sharing points is 1, and data quality is 1. Based on the findings, we conclude that when data are highly sensitive and its quality is good, the number of the sharing points interested in that particular data increases in the system.

Accordingly, we define 7.5 as a threshold measurement to consider data requests with severe data sensitivity, regular sharing points, and standard data quality as the ones to be allocated within the blockchain embedded in cloud and fog environments. If the calculated RoA value of a data request is below this threshold, it is automatically stored off-chain to keep a historical record of IoT data transactions. Although our approach proposes 7.5 as a threshold value to support on-chain data allocation decisions, it can be changed based on the system administrator and IoT system requirements.

Fig. 7. FogBus sleep apnea analysis prototype [16].

VII. PERFORMANCE EVALUATION

In this section, we instantiate the data allocation mechanism in two commonly used IoT-blockchain architectural styles for the healthcare study. Next, we measure the efficiency of the data allocation mechanism in the two architectures (i.e., blockchain-based fog and cloud) in terms of latency, network usage, and energy consumption.

A. Evaluation Goals

We summarize the motivations for the integration of the data allocation mechanism in the blockchain-based fog and cloud architectures as follows.

1) Assess the effectiveness of the data allocation mechanism by either enabling or disabling it in the two IoT-blockchain architectural styles. It ensures flexibility in the system and satisfies the end users and service providers requirements.
2) Evaluate the performance of the refined blockchain-based cloud and fog architectures in terms of energy consumption, latency, blockchain size, and network usage. It compares the performance of the two architectural styles when a huge number of requests are generated simultaneously.

B. Simulation Environment

The evaluation of the data allocation mechanism consists of two stages: 1) collect, process, and store context parameters of each IoT data request using FogBus and 2) simulate the data controller using MATLAB to calculate the RoA value and support data allocation.

FogBus is a real-world lightweight-blockchain framework that integrates IoT, fog, edge, cloud, and blockchain. Fig. 7 shows the FogBus-enabled sleep apnea analysis prototype presented in [16].

Table IV includes the simulation parameters. The setup of the hardware components and their configuration are given as follows.

1) *IoT Device:* Pulse oximeter, 1.5 V, Bluetooth 4.0, UFT-8 data encoding.

TABLE IV
SIMULATION PARAMETERS

Parameters	Values
Analysis task:	
Interval between the creation of consecutive data processing requests	5 seconds
Data recording time per processing requests	3 minutes
Pulse oximeter:	
Pulse oximeter signal length	18 KB
Sensing frequency	2 signal per second
WLAN:	
Download speed	7 Mbps
Upload speed	2 Mbps

2) *Gateway Node:* Oppo A77T smartphone, Android 7.1.1.
3) *Master Node:* Dell Latitude D630 Laptop, Intel Core 2 Duo CPU E6550 @ 2.33-GHz 2-GB DDR2 RAM, 32-b, Windows 7, Apache Server 2.4.34, Java 1.6, MySQL 5.6, .NET 3.5, and Aneka 3.1.
4) *Worker Node:* Raspberry Pi 3, ARM Cortex A53 quad-core SoC CPU@ 1.4-GHz 1-GB LPDDR2 SDRAM, Raspian Stretch, Apache Server 2.4.34, Java 1.6, and MySQL 5.6.
5) *Cloud:* Microsoft Azure B1s Machine, 1vCPU, 1-GB RAM, 2-GB SSD, Windows Server 2010, .NET 3.5, and Aneka 3.1.

Initially, the oximeter collects timestamp, heartbeat, and blood oxygen level for 1 h of sleep study and transmits it to the gateways in the data controller tier which keeps an internal list of records. Once the recordings terminate, the data controller receives the oximeter data along with the number of sharing points interested in the data (i.e., doctors, hospitals, laboratories, pharmacies, etc.), and the device accuracy measurement. Here, we use MATLAB to design and simulate the data controller with its corresponding inputs, membership functions, and fuzzy rules. We calculate the RoA value of each data request to realize the data allocation mechanism in the blockchain-based cloud and fog architectures using the FogBus framework.

We define the following concrete metrics to evaluate the efficiency of the data allocation mechanism in the blockchain-based fog and cloud architectures.

1) *Size of Blockchain:* Average size of the blockchain in the broker node and cloud VMs.
2) *Average Latency:* Data access latency to retrieve data from fog nodes (i.e., broker node and worker nodes) and cloud VM. Since we use FogBus framework [16] as our simulation environment, it can directly provide the measures of the latency. In particular, its measured latency refers to the overall system latency (i.e., data processing, network propagation delay, and OS delay).
3) *Network Usage:* The load in the network when the data allocation mechanism is deployed in the blockchain-based fog and cloud.
4) *Energy Consumption:* The average energy usage of the broker node for the blockchain-based fog and the average energy usage of the Azure VM for the blockchain-based cloud to support the blockchain. In particular, we monitored the energy consumption in the broker node and Azure VM via enabling/disabling the data allocation mechanism by a Joulemeter [52] which can estimate the energy consumption of runtime applications.

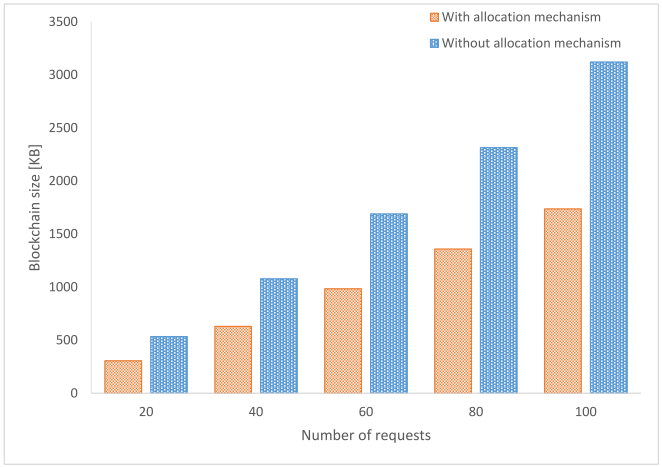


Fig. 8. Size of Blockchain in kilobytes—with/without the allocation mechanism.

C. Result Analysis

We describe the performance results of refined blockchain-based cloud and fog architectures.

1) Size of Blockchain: Fig. 8 depicts the estimated storage size in kB of the blockchain-based with/without the realization of the data allocation mechanism. Overall, the blockchain size increases linearly when the mechanism is not applied in the architectural styles since all IoT data requests are allocated within the blockchain embedded in fog and cloud without considering its limited storage capacity. On the contrary, the implementation of the data allocation mechanism in the refined IoT-blockchain architectural styles ensures a reduction of around 42% on average in blockchain size. In fact, when 100 data requests are executed in the system, the storage size of a blockchain decreases around 44% on average, and about 42% when 20 data requests are carried out into the refined blockchain-based cloud and fog, respectively. From the results, we conclude that the implementation of the data allocation mechanism alleviates the storage capacity of the blockchain since only data request with a high RoA value is stored within the blockchain-based cloud and fog architectures.

2) Average Latency: Fig. 9 illustrates the service delivery latency in seconds when the data allocation mechanism is enabled and disabled into the blockchain-based (a) cloud and (b) fog architectures. To simulate the propagation delay from the cloud, we first connect the broker node and Azure VM through a virtual network of 4 Mb/s. In particular, we define the propagation delay in the blockchain-based fog architecture as the delay from the broker node to the worker nodes. Similarly, we measure the propagation delay in the blockchain-based cloud architecture as the delay from the broker node to the cloud VM. The results show that the network propagation delay in the blockchain-based cloud is almost two times more than in the blockchain-based fog when the data allocation

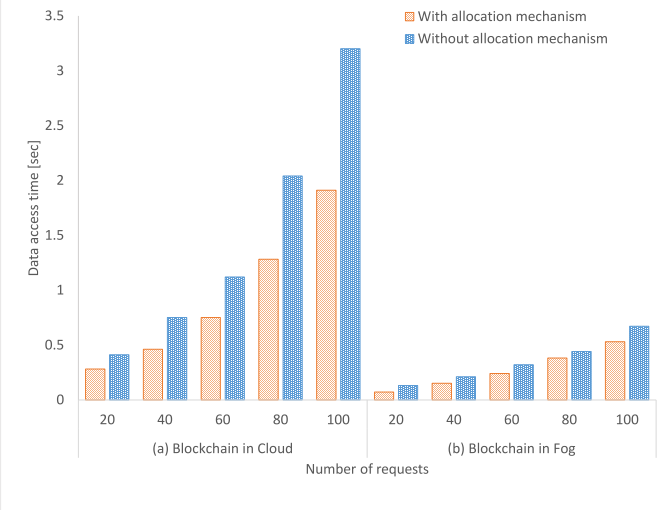


Fig. 9. Data access time in seconds—with/without the allocation mechanism in (a) blockchain-based cloud and (b) blockchain-based fog.

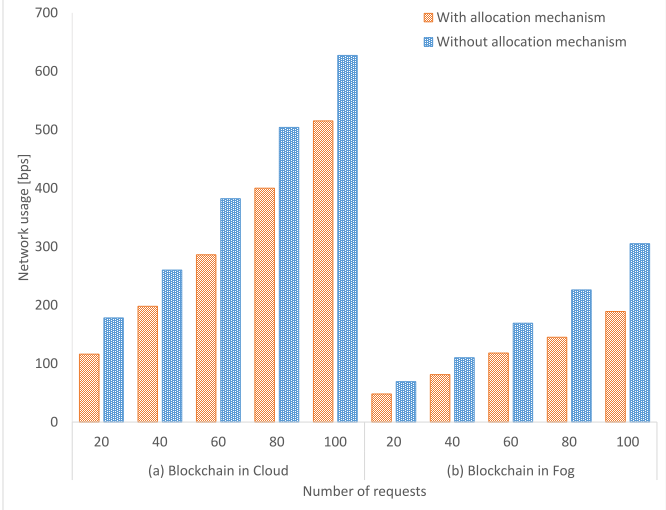


Fig. 11. Network usage in bits per second—with/without the allocation mechanism in (a) blockchain-based cloud and (b) blockchain-based fog.

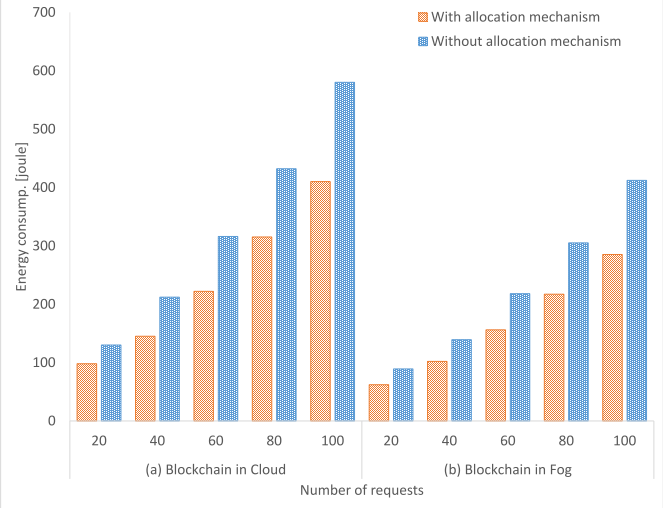


Fig. 10. Energy consumption in Joule—with/without the allocation mechanism in (a) blockchain-based cloud and (b) blockchain-based fog.

mechanism is not executed. However, its implementation in the blockchain-based cloud contributes to a latency reduction of 36% on average and about 27% in the blockchain-based fog. These results show that the data allocation mechanism effectively reduces the amount of data to be sent to the blockchain whether it is executed in the blockchain-cloud and fog architectures. In addition, as the size of the data chunk to be recorded in the blockchain is not huge, the latency will not differ significantly between the two architectures. Thus, we conclude that the service deliver latency mainly depends on the network propagation delay, which is low in the blockchain-fog architecture since fog nodes are located in single-hop proximity to where data are collected.

3) Energy Consumption: Fig. 10 shows the estimated amount of energy consumption in joules when the data allocation mechanism is enabled and disabled in the blockchain-based (a) cloud and (b) fog architectures. Here, the energy consumption in the blockchain-based cloud is about 32% more than in the blockchain-based fog when the mechanism is not integrated into the system. These results evidence that the tasks performed in the cloud are complex and require additional computing, storage, and networking resources. On the other hand, the realization of the data allocation mechanism leads to an energy reduction of 28% on average in the blockchain-based cloud and fog architectures. When 100 data requests are executed in the system, the energy consumption reaches 410 J in the refined blockchain-based cloud and just above 285 in the refined blockchain-based fog. In other words, when a high number of data requests are allocated in the blockchain-based fog, the energy consumption is lower compared to the cloud since fog devices are located in single-hop proximity of IoT devices. On the contrary, when a high number of data requests are allocated in the blockchain-based cloud, the energy consumption is the same or higher than in the fog since cloud servers are located in multihop proximity of IoT devices. From these observations, we conclude that the refined blockchain-based fog saves more energy compared to the refined blockchain-based cloud since IoT data requests are processed close to where the data are collected.

4) Network Usage: Fig. 11 illustrates the network usage in bits per second when the data allocation mechanism is enabled and disabled in the blockchain-based (a) cloud and (b) fog architectures. The figure shows that the network usage in the blockchain-based cloud is about two times more than in the blockchain-based fog when the data allocation mechanism is not integrated into the system. These results show that the fog outperforms the cloud since it enables local networking resources to handle the IoT data requests. However, the integration of the data allocation mechanism reduces network usage in the blockchain-based cloud and fog architectures. Approximately, 32% of network usage is reduced in the refined blockchain-based fog and just above 24% on average in the refined blockchain-based cloud. In fact, the network usage in refined blockchain-based fog is about 189 b/s, while it reaches a peak over 515 b/s in the refined blockchain-based cloud. Although the implementation of the blockchain

TABLE V
FUZZY LOGIC DECISION-MAKING RESULTS COMPARED WITH STATE-OF-THE-ART APPROACHES

Approaches	CPU usage	Execution time
Logistic regression	84%	0.239 sec
Decision tree	51%	0.156 sec
Fuzzy logic	30%	0.112 sec

in fog and cloud environments increases network usage due to the security mechanisms implemented (e.g., encryption algorithms), blockchain-based fog reports less network usage than blockchain-based cloud since fog uses local networking resource which reduces latency and bandwidth consumption in the network.

D. Performance Comparison With Alternative Decision-Making Mechanisms

We further compare the effectiveness of our approach against the existing alternative decision-making mechanisms for data management. Specifically, we first survey the literature and identify that logistic regression [18], [38] and decision tree [20], [21], [53] are often considered as alternative decision-making approaches. Next, we conduct experiments for performance comparison with such approaches to show the benefit of our technique. Our experiment results suggest that our approach incurs a significantly declining CPU usage and overall execution time in the broker node as described in Table V. In particular, the CPU usage in the broker node achieves 84% and 51% when running logistic regression and decision tree, respectively, while our approach can reduce the CPU usage to 30%. Similarly, our approach (0.112 s) can outperform all the compared approaches (0.156 s, 0.239 s) in terms of execution time.

VIII. RELATED WORK

In this section, we briefly summarize a subset of relevant work to our system.

A. Fuzzy Logic in IoT

There exist several applications of fuzzy logic in IoT systems as described below. Vani and Neeralagi [34] proposed a real-time IoT health monitoring system for elderly people that collects environmental data and uses fuzzy logic to simplify its interpretation and take decisions accordingly. Santamaria *et al.* [54] proposed a fuzzy-logic approach that learns customer habits through body sensors and discovers outliers warning signals to minimize the risk of false alarms. Similarly, Bhunia *et al.* [55] proposed a healthcare system based on fuzzy logic for a smart city where sensor data are collected (i.e., SPO2, ECG, airflow, temperature, etc.) to support decision making about true conditions of the patient, e.g., weak heart, shock, and respiratory problem. In addition to the healthcare domain, Meana-Llorián *et al.* [41] designed a fuzzy-logic system that autonomous controls indoor temperature using external climate conditions that results in 40% energy reduction. Novilla *et al.* [56] designed a manufacturing monitoring system based on fuzzy logic that uses temperature and smoke sensors to capture normal conditions of the manufacturing machines and build a reference model to inform the machine health status and provide accurate failure predictions. Mahalle *et al.* [57] presented a fuzzy-logic approach to enhance trust-based access control in IoT that use vague values of Experience (EX), Knowledge (KN), and Recommendation (RC) to authorize devices in the IoT network. Another approach presents a fuzzy logic framework to determine employee performance appraisal based on the IoT data [58]. Globa *et al.* [59] proposed the use of a fuzzy-logic mechanism for big data processing in IoT networks in order to improve the performance and reduce computational costs of complex machine-learning algorithms.

B. Decision-Making Mechanisms in IoT Systems

There exists a considerable body of literature on decision-making mechanisms applied on different domains, such as healthcare, manufacturing, and control systems [38]. Lowe and Parvar [18] proposed a logistic regression approach for decision-making about bid/no-bid from contractors in a construction company. Similarly, Young *et al.* [19] used a regression model to predict diabetes severity index and risk of mortality. Ohno-Machado *et al.* [20] proposed the use of decision tree and fuzzy logic as the decision models to select the optimal vaccination strategy. López-Vallverdú *et al.* [21] presented a decision model based on a decision tree algorithm that combines relevant healthcare criteria for screening and diagnosis. Also, Chern *et al.* [53] proposed a decision tree model that delivers optimal telehealth services and reduces the misuse of resources. Karan *et al.* [60] presented a diagnostic illness system based on artificial neural networks that collect data from small mobile devices. Similarly, Burke *et al.* [61] relied on artificial neural networks for improving the accuracy of cancer prediction. Also, Ting *et al.* [62] proposed a diagnostic system for obstructive sleep apnea based on decision tree algorithms, which are able to perform automatic feature selection. Timuş and Bolat *et al.* [63] presented a k-nearest neighbors (k-NN) classifier for determining the sleep apnea types. The fuzzy decision tree is proposed for classification and prediction problems [64], [65].

Our proposed approach differs from the aforementioned works since we have considered the use of fuzzy logic for deriving on-chain allocation decisions based on multiple context parameters. Specifically, we design a data controller based on fuzzy logic that handles multiple context parameters, e.g., data, network, and quality to calculate the RoA value of each IoT data request and support on-chain data allocation. The RoA value is used as a threshold measurement to decide which data request needs to be allocated within the blockchain or stored off-chain, e.g., cloud database. Moreover, the realization of the data allocation mechanism in the two IoT-blockchain architectural styles leads to refinements that are analyzed from an abstract level by proposing a four-tier abstraction, i.e., the IoT tier, the data controller tier, the fog tier, and the cloud tier. To demonstrate the effectiveness of our approach, we instantiate it in the blockchain-based cloud and fog architectures and evaluate their performance in terms of network usage, latency, energy consumption, and blockchain storage.

IX. Discussion

We have shown the realization of the data allocation mechanism in two commonly used architectures across the implementations that integrate blockchain in IoT systems, i.e., blockchain-based cloud and fog [15], [17]. Although we have chosen these architectural styles as a way to illustrate our approach, it can continue to work in other styles. In particular, an alternative data management strategy could be having the blockchain as a separate network in fog or cloud environments where each block stores only the hash of the data and data address of the relevant data while maintaining the raw data in the cloud to meet IoT system requirements. Moreover, we can develop a market-based mechanism to decide the utility improvement of using fog and cloud environments or secure platforms (i.e., blockchain-based cloud and fog) for IoT data allocation considering the cost, QoS requirements, and constraints imposed by each hosting environment. In this model, users or service providers can be charged based on a pay-as-you-go or subscription fee to decide when to use the normal fog and cloud or secure environments (i.e., blockchain-based cloud and fog) for storing IoT data.

X. Conclusion

In this article, we identify a number of architectural significant requirements for developing a data-centric approach that supports data allocation in IoT systems supported with blockchain. To meet these requirements, we propose a novel context-aware data allocation mechanism that calculates the RoA value of each IoT data request based on multiple context parameters to decide its on-chain allocation. The mechanism relies on the design of a data controller based on fuzzy logic that extracts context parameters of each data request, e.g., data, network, and quality to determine the RoA value which is used as a threshold measurement to decide which data request needs to be stored within the blockchain or allocated off-chain, i.e., cloud database. Moreover, we enrich the two commonly used IoT-blockchain architectural styles supported by fog and cloud with the data allocation mechanism where we introduce a data controller tier between the IoT tier and the fog tier to handle on-chain allocation decisions in real time. To evaluate the effectiveness of our approach, we instantiate the blockchain-based cloud and fog architectures in a healthcare example using the FogBus framework. We conduct several experiments to measure the latency, energy consumption, network usage, and blockchain storage in refined architectural styles. The performance evaluation suggests that latency is reduced by 36% in the refined blockchain-based cloud and about 27% in the refined blockchain-based fog. Similarly, energy consumption is reduced by averagely 28% in the refined blockchain-based cloud and fog. Moreover, the network usage is reduced by 32% in the refined blockchain-based fog and 24% in the refined blockchain-based cloud.

Although our approach can be best described as a reactive data allocation mechanism that continuously retrieves context parameters for IoT data requests to calculate the RoA value, it can be developed as a dynamic and adaptive controller. To this end, one of our future work direction is investigating how self-adaptive mechanisms and AI, in particular genetic algorithms, can be applied to develop an intelligent controller that handles environmental and internal uncertainties and derives optimal on-chain data allocation decisions [66], [67]. Specifically, we aim to design a data controller that continuously learns from previous executions and its environment to improve the current on-chain data allocation mechanism. Furthermore, we aim to investigate the generality of the proposed mechanism and its application on other alternative styles taking advantage of private and public blockchains. We also aim to evaluate the performance of the refined IoT-blockchain architectural styles in a real environment by integrating other sensors (e.g., temperature and air quality) to improve the calculation of the RoA value in order to minimize the risk of uncertainty in data allocation decisions.

References

[1] I.-L. Yen, F. Bastani, W. Zhu, H. Moeini, S. Hwang, and Y. Zhang, "Service-oriented IoT modeling and its deviation from software services," in *Proc. IEEE Symp. Service Orient. Syst. Eng. (SOSE)*, Mar. 2018, pp. 40–47.

[2] I.-L. Yen, S. Zhang, F. B. Bastani, and Y. Zhang, "A framework for IoT-based monitoring and diagnosis of manufacturing systems," in *Proc. IEEE Symp. Service Orient. Syst. Eng. (SOSE)*, San Francisco, CA, USA, Apr. 2017, pp. 1–8. [Online]. Available: https://doi.org/10.1109/SOSE.2017.26

[3] *Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units by 2020*, Gartner, Stamford, CT, USA, 2013.

[4] D. Evans, "The Internet of Things: How the next evolution of the Internet is changing everything," San Jose, CA, USA, CISCO, White Paper, pp. 1–11, 2011.

[5] L. Fan, J. R. Gil-Garcia, D. Werthmuller, G. B. Burke, and X. Hong, "Investigating blockchain as a data management tool for IoT devices in smart city initiatives," in *Proc. ACM 19th Annu. Int. Conf. Digit. Govt. Res. Governance Data Age*, 2018, p. 100.

[6] M. A. Razzaque, M. Milojevic-Jevric, A. Palade, and S. Clarke, "Middleware for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 3, no. 1, pp. 70–95, Feb. 2016.

[7] T. Zheng *et al.*, "SmartVM: A SLA-aware microservice deployment framework," *World Wide Web*, vol. 22, no. 1, pp. 275–293, 2019. [Online]. Available: https://doi.org/10.1007/s11280-018-0562-5

[8] M. Zhang, Y. Zhang, L. Zhang, C. Liu, and S. Khurshid, "DeepRoad: GAN-based metamorphic testing and input validation framework for autonomous driving systems," in *Proc. 33rd ACM/IEEE Int. Conf. Autom. Softw. Eng. (ASE)*, Montpellier, France, Sep. 2018, pp. 132–142. [Online]. Available: https://doi.org/10.1145/3238147.3238187

[9] D. Yu, Y. Jin, Y. Zhang, and X. Zheng, "A survey on security issues in services communication of microservices-enabled fog applications," *Concurrency Comput. Pract. Exp.*, vol. 31, no. 22, 2019, Art. no. e4436. [Online]. Available: https://doi.org/10.1002/cpe.4436

[10] A. Botta, W. De Donato, V. Persico, and A. Pescapé, "Integration of cloud computing and Internet of Things: A survey," *Future Gener. Comput. Syst.*, vol. 56, pp. 684–700, Mar. 2016.

[11] X. Xu *et al.*, "The blockchain as a software connector," in *Proc. 13th Working IEEE/IFIP Conf. Softw. Archit. (WICSA)*, 2016, pp. 182–191.

[12] P. Brody and V. Pureswaran, *Device Democracy: Saving the Future of the Internet of Things*, IBM, Armonk, NY, USA, Sep. 2014.

[13] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *Proc. IEEE Int. Congr. Big Data (BigData Congr.)*, 2017, pp. 557–564.

[14] T. M. Fernández-Caramés and P. Fraga-Lamas, "A review on the use of blockchain for the Internet of Things," *IEEE Access*, vol. 6, pp. 32979–33001, 2018.

[15] M. Samaniego and R. Deters, "Blockchain as a service for IoT," in *Proc. IEEE Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber Phys. Soc. Comput. (CPSCom) IEEE Smart Data (SmartData)*, 2016, pp. 433–436.

[16] S. Tuli, R. Mahmud, S. Tuli, and R. Buyya, "FogBus: A blockchain-based lightweight framework for edge and fog computing," 2018. [Online]. Available: arXiv:1811.11978.

[17] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *Proc. 17th IEEE/ACM Int. Symp. Clust. Cloud Grid Comput.*, 2017, pp. 468–477.

[18] D. J. Lowe and J. Parvar, "A logistic regression approach to modeling the contractor's decision to bid," *Construction Manag. Econ.*, vol. 22, no. 6, pp. 643–653, 2004.

[19] B. A. Young *et al.*, "Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization," *Amer. J. Manag. Care*, vol. 14, no. 1, p. 15, 2008.

[20] L. Ohno-Machado, R. Lacson, and E. Massad, "Decision trees and fuzzy logic: A comparison of models for the selection of measles vaccination strategies in Brazil," in *Proc. AMIA Symp.*, 2000, p. 625.

[21] J. A. López-Vallverdú, D. RiañO, and J. A. Bohada, "Improving medical decision trees by combining relevant health-care criteria," *Expert Syst. Appl.*, vol. 39, no. 14, pp. 11782–11791, 2012.

[22] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, 2010.

[23] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.

[24] S. Hammoudi, Z. Aliouat, and S. Harous, "Challenges and research directions for Internet of Things," *Telecommun. Syst.*, vol. 67, no. 2, pp. 367–385, 2018.

[25] M. Atzori, "Blockchain-based architectures for the Internet of Things: A survey," 2017.

[26] F. Computing, "The Internet of Things: Extend the cloud to where the things are," San Jose, CA, USA, Cisco, White Paper, 2015.

[27] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: Towards a comprehensive definition of fog computing," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 5, pp. 27–32, 2014.

[28] J. Ni, K. Zhang, X. Lin, and X. S. Shen, "Securing fog computing for Internet of Things applications: Challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 601–628, 1st Quart., 2017.

[29] A. Stanciu, "Blockchain based distributed control system for edge computing," in *Proc. 21st Int. Conf. Control Syst. Comput. Sci. (CSCS)*, 2017, pp. 667–671.

[30] M. Yannuzzi, R. Milito, R. Serral-Gracià, D. Montero, and M. Nemirovsky, "Key ingredients in an IoT recipe: Fog computing, cloud computing, and more fog computing," in *Proc. IEEE 19th Int. Workshop Comput.-Aided Model. Design Commun. Links Netw. (CAMAD)*, 2014, pp. 325–329.

[31] X. Masip-Bruin, E. Marín-Tordera, G. Tashakor, A. Jukan, and G.-J. Ren, "Foggy clouds and cloudy fogs: A real need for coordinated management of fog-to-cloud computing systems," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 120–128, Oct. 2016.

[32] X. Xu *et al.*, "A taxonomy of blockchain-based systems for architecture design," in *Proc. IEEE Int. Conf. Softw. Archit. (ICSA)*, 2017, pp. 243–252.

[33] J. Eberhardt and S. Tai, "On or off the blockchain? Insights on off-chaining computation and data," in *Proc. Eur. Conf. Service Orient. Cloud Comput.*, 2017, pp. 3–15.

[34] K. Vani and R. R. Neeralagi, "IoT based health monitoring using fuzzy logic," *Int. J. Comput. Intell. Res.*, vol. 13, no. 10, pp. 2419–2429, 2017.

[35] J. Durán, S. Esnaola, R. Rubio, and Á. Iztueta, "Obstructive sleep Apnea–Hypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 yr," *Amer. J. Res. Critical Care Med.*, vol. 163, no. 3, pp. 685–689, 2001.

[36] D. Weyns, G. S. Ramachandran, and R. K. Singh, "Self-managing Internet of Things," in *Proc. Int. Conf. Current Trends Theory Practice Informat.*, 2018, pp. 67–84.

[37] M. U. Iftikhar, G. S. Ramachandran, P. Bollansée, D. Weyns, and D. Hughes, "DeltaIoT: A self-adaptive Internet of Things exemplar," in *Proc. 12th Int. Symp. Softw. Eng. Adapt. Self Manag. Syst.*, 2017, pp. 76–82.

[38] G. Gürsel *et al.*, "Healthcare, uncertainty, and fuzzy logic," *Digit. Med.*, vol. 2, no. 3, p. 101, 2016.

[39] A. Magruk, "The most important aspects of uncertainty in the Internet of Things field–context of smart buildings," *Procedia Eng.*, vol. 122, pp. 220–227, Oct. 2015.

[40] G. Cueva-Fernandez, J. P. Espada, V. García-Díaz, and R. Gonzalez-Crespo, "Fuzzy decision method to improve the information exchange in a vehicle sensor tracking system," *Appl. Soft Comput.*, vol. 35, pp. 708–716, Oct. 2015.

[41] D. Meana-Llorián, C. G. García, B. C. P. G-bustelo, J. M. C. Lovelle, and N. Garcia-Fernandez, "IoFClime: The fuzzy logic and the Internet of Things to control indoor temperature regarding the outdoor ambient conditions," *Future Gener. Comput. Syst.*, vol. 76, pp. 275–284, Nov. 2017.

[42] A. Patel and T. A. Champaneria, "Fuzzy logic based algorithm for context awareness in IoT for smart home environment," in *Proc. IEEE Region 10 Conf. (TENCON)*, 2016, pp. 1057–1060.

[43] I. Makhdoom, M. Abolhasan, H. Abbas, and W. Ni, "Blockchain's adoption in IoT: The challenges, and a way forward," *J. Netw. Comput. Appl.*, vol. 125, pp. 251–279, Jan. 2019.

[44] A. Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, and F. Wang, "Secure and trustable electronic medical records sharing using blockchain," in *Proc. AMIA Annu. Symp.*, vol. 2017, 2017, p. 650.

[45] Y. Bai and D. Wang, "Fundamentals of fuzzy logic control—Fuzzy sets, fuzzy rules and defuzzifications," in *Advanced Fuzzy Logic Technologies in Industrial Applications*. London, U.K.: Springer, 2006, pp. 17–36.

[46] A. Ansari and A. A. Bakar, "A comparative study of three artificial intelligence techniques: Genetic algorithm, neural network, and fuzzy logic, on scheduling problem," in *Proc. IEEE 4th Int. Conf. Artif. Intell. Appl. Eng. Technol. (ICAIET)*, 2014, pp. 31–36.

[47] R. Mahmud, S. N. Srirama, K. Ramamohanarao, and R. Buyya, "Quality of experience (QoE)-aware placement of applications in fog computing environments," *J. Parallel Distrib. Comput.*, vol. 132, pp. 190–203, Oct. 2019.

[48] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man Mach. Stud.*, vol. 7, no. 1, pp. 1–13, 1975.

[49] M. E. F. Morán and N. A. P. Viera, "Comparative study for DC motor position controllers," in *Proc. IEEE Ecuador Tech. Meeting (ETCM)*, 2017, pp. 1–6.

[50] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Proc. Int. Symp. Handheld Ubiquitous Comput.*, 1999, pp. 304–307.

[51] S. P. Athan and J. E. Scharf, "Portable pulse Oximeter," U.S. Patent 5 575 284, Nov. 19, 1996.

[52] *JouleMeter: Computational Energy Measurement and Optimization*. Accessed: Nov. 20, 2019. [Online]. Available: https://www.microsoft.com/en-us/research/project/joulemetercomputational-energy-measurement-and-optimization

[53] C.-C. Chern, Y.-J. Chen, and B. Hsiao, "Decision tree-based classifier in providing telehealth service," *BMC Med. Informat. Decision Making*, vol. 19, no. 1, p. 104, 2019.

[54] A. F. Santamaria, P. Raimondo, F. De Rango, and A. Serianni, "A two stages fuzzy logic approach for Internet of Things (IoT) wearable devices," in *Proc. IEEE 27th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2016, pp. 1–6.

[55] S. S. Bhunia, S. K. Dhar, and N. Mukherjee, "iHealth: A fuzzy approach for provisioning intelligent health-care system in smart city," in *Proc. IEEE 10th Int. Conf. Wireless Mobile Comput. Netw. Commun. (WiMob)*, 2014, pp. 187–193.

[56] A. G. P. Novilla, A. A. N. Balute, and D. B. Gonzales, "The use of fuzzy logic for online monitoring of manufacturing machine: An intelligent system," *Circulation Comput. Sci.*, vol. 2, pp. 31–39, 2017.

[57] P. N. Mahalle, P. A. Thakre, N. R. Prasad, and R. Prasad, "A fuzzy approach to trust based access control in Internet of Things," in *Proc. IEEE Wireless VITAE*, 2013, pp. 1–5.

[58] J. Kaur and K. Kaur, "A fuzzy approach for an IoT-based automated employee performance appraisal," *Comput. Mater. Continua*, vol. 53, no. 1, pp. 24–38, 2017.

[59] L. Globa, V. Kurdecha, I. Ishchenko, A. Zakharchuk, and N. Kunieva, "Fuzzy logic usage for the data processing in the Internet of Things networks," 2018.

[60] O. Karan, C. Bayraktar, H. Gümüşkaya, and B. Karlık, "Diagnosing diabetes using neural networks on small mobile devices," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 54–60, 2012.

[61] H. B. Burke *et al.*, "Artificial neural networks improve the accuracy of cancer survival prediction," *Cancer*, vol. 79, no. 4, pp. 857–862, 1997.

[62] H. Ting, Y.-T. Mai, H.-C. Hsu, H.-C. Wu, and M.-H. Tseng, "Decision tree based diagnostic system for moderate to severe obstructive sleep apnea," *J. Med. Syst.*, vol. 38, no. 9, p. 94, 2014.

[63] O. H. Timuş and E. D. Bolat, "k-NN-based classification of sleep Apnea types using ECG," *Turkish J. Elect. Eng. Comput. Sci.*, vol. 25, no. 4, pp. 3008–3023, 2017.

[64] J. F. Baldwin and D. W. Xie, "Simple fuzzy logic rules based on fuzzy decision tree for classification and prediction problem," in *Proc. Int. Conf. Intell. Inf.*, 2004, pp. 175–184.
[65] C.-Y. Fan, P.-C. Chang, J.-J. Lin, and J. Hsieh, "A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 632–644, 2011.
[66] R. Calinescu, D. Weyns, S. Gerasimou, M. U. Iftikhar, I. Habli, and T. Kelly, "Engineering trustworthy self-adaptive software with dynamic assurance cases," *IEEE Trans. Softw. Eng.*, vol. 44, no. 11, pp. 1039–1069, Nov. 2018.
[67] E. Flores-Morán, W. Yánez-Pazmiño, and J. Barzola-Monteses, "Genetic algorithm and fuzzy self-tuning PID for DC motor position controllers," in *Proc. 19th Int. Carpathian Control Conf. (ICCC)*, 2018, pp. 162–168.

Wendy Yánez received the master's degree (with Distinction) in computer security from the University of Birmingham, Birmingham, U.K. She is currently pursuing the Ph.D. degree in the Join-Program University of Birmingham and Southern University of Science and Technology. She was awarded with the Scholarship for this collaborative program.

She is also collaborating with the Facultad de Ingeniería en Electricidad y Computación, Escuela Superior Politécnica del Litoral, ESPOL, ESPOL Polythecnic University, Guayaquil, Ecuador. Her research interests include Internet of Things, fog computing, and blockchain.

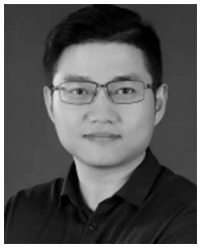
Redowan Mahmud received the B.Sc. degree from the Department of Computer Science and Engineering, University of Dhaka, Dhaka, Bangladesh, in 2015. He is currently pursuing the Ph.D. degree with the Cloud Computing and Distributed Systems Laboratory, Department of Computing and Information Systems, University of Melbourne, Melbourne VIC, Australia.

His research interests include Internet of Things, fog computing, and mobile cloud computing.

Mr. Mahmud was awarded the Melbourne International Research Scholarship and the Melbourne International Fee Remission Scholarship supporting his studies.

Rami Bahsoon (Member, IEEE) received the Ph.D. degree in software engineering from University College London, London, U.K., for his research on evaluating software architecture stability using real options, and he attended London Business School for MBA-level studies in technology strategy and dynamics.

He is a Senior Lecturer of software engineering (Associate Professor) and leads the software engineering for/in the Cloud Interest Group, University of Birmingham, Birmingham, U.K. The group's research aims at developing architecture and frameworks to support and reason about dependable complex software systems, where the investigations span cloud computing architectures and their economics. He published extensively in the area of economics-driven software engineering, cloud software engineering, and utility computing and co-edited a book *Software Architecture and Software Quality* and a book *Economics-Driven Software Architecture* (Elsevier).



Yuqun Zhang (Member, IEEE) received the B.S. degree from Tianjin University, Tianjin, China, the M.S. degree from the University of Rochester, Rochester, NY, USA, and the Ph.D. degree from the University of Texas at Austin, Austin, TX, USA.

He is an Assistant Professor with the Southern University of Science and Technology, Shenzhen, China. His research interests include software engineering and services computing.

Rajkumar Buyya received the Ph.D. degree in computer science and software engineering from Monash University, Melbourne, VIC, Australia, in 2002.

He is a Professor and a Future Fellow of the Australian Research Council, and the Director of the Cloud Computing and Distributed Systems Laboratory, University of Melbourne, Melbourne, VIC, Australia. He is also serving as the Founding CEO of Manjrasoft, a spinoff company of the University, commercializing its innovations in Cloud Computing. He has authored more than 425 publications and four textbooks, including *Mastering Cloud Computing* published by McGraw Hill and Elsevier/Morgan Kaufmann, 2013, for Indian and international markets, respectively. He is one of the highly cited authors in computer science and software engineering worldwide. Microsoft Academic Search Index ranked him as the world's top author in distributed and parallel computing from 2007 to 2012. Software technologies for grid and cloud computing developed under his leadership have gained rapid acceptance and are in use at several academic institutions and commercial enterprizes in 40 countries around the world. He has led the establishment and development of key community activities, including serving as the Foundation Chair of the IEEE Technical Committee on Scalable Computing and five IEEE/ACM conferences.

Prof. Buyya received the "2009 IEEE Medal for Excellence in Scalable Computing" from the IEEE Computer Society. Manjrasoft's Aneka Cloud technology developed under his leadership has received the "2010 Frost & Sullivan New Product Innovation Award" and the "2011 Telstra Innovation Challenge, People's Choice Award." He served as the Founding Editor-in-Chief for the IEEE TRANSACTIONS ON CLOUD COMPUTING. He is currently serving as the Co-Editor-in-Chief for the *Journal of Software: Practice and Experience*, which was established over 45 years ago.