Addressing the Future Data Management Challenges in IoT: A Proposed Framework

Mohammad Asad Abbasi[1], Zulfiqar A. Memon[1], Tahir Q. Syed[1], Jamshed Memon[2], Rabah Alshboul[3]
[1]Department of Computer Science, National University of Computer & Emerging Sciences, Karachi, Pakistan
[2]Department of Computer Science, Barrett Hodgson University, Karachi, Pakistan
[3] Department of Computer Science, Al Al-Bayt University, Mafraq, Jordan

***Abstract*—Internet of Thing (IoT) has been attracting the interest of researchers in recent years. Traditionally, only handful types of devices had the capability to be connected to internet/intranet, but due to the latest developments in RFID, NFC, smart sensors and communication protocols billions of heterogeneous devices are being connected each year. From smart phones uploading the data regarding location and fitness to smart grids uploading the data regarding energy consumption and distribution, these devices are generating a huge amount of data each passing moment. This research paper proposes a data management framework to securely manage the huge amount of data that is being generated by IoT enabled devices. The proposed framework is divided into nine layers. The framework incorporates layers such as data collection layer, fog computing layer, integrity management layer, security layer, data aggregation layer, data analysis layer, data storage layer, application layer and archiving layer. The security layer has been proposed as a background layer because all layers shall ensure the privacy and security of the data. These layers will help in managing the data from the point where it is generated by an IoT enabled device until the point where the data is archived at the data center.**

Keywords—IoT; Data Management; Cloud Computing; Big Data; Smart Devices; Interoperability; Privacy; Trust

I. INTRODUCTION

Internet of Things is one of the concepts, which tends to build a new future of computing by taking every smart object into a globally connected network capable of sensing, communicating, information sharing and performing smart analytics for different applications [1][7]. This is the result of rising technological evolution of computing devices and its use in different sectors like healthcare, automotive, education and sports. The excessive use of smart objects in human life has pushed the researchers towards the design and development of tools and techniques that can connect these smart devices to a global network. Emphasis has been to enhance the efficiency of these smart devices to generate less, but meaningful data that can be efficiently transported and analysed on a cloud before being stored. Last decade is a witness of the development of different network protocols, computing devices and storage devices that have helped in the rapid deployment of IoT enabled devices. [1][5][7][8][13].

Furthermore, it has been observed that this wave of smart devices is serving in different areas such as education, medical, military, research, sports and industries [5][15][17]. One of the application switches that IoT has made possible is a smart home concept. Smart home offers services like access control, home monitoring, safety and central control of numerous home appliances to its owner [4][11][15]. The basic idea of smart homes is to connect home appliances to network and employ the use of some standard protocols for communications. Smart sensors and cameras are utilised for this purpose [5] [15]. Another application that can be witnessed is smart agriculture where IoT exploits smart sensors and RFIDs to change the shape of traditional decision making regarding crops. IoT has enabled the farmers to be aware of information related to different field parameters like humidity, moisture, temperature and wind speed. This makes it possible for farmers to take timely and more accurate decisions for enhancing crop productivity and quality.

One more key application area is supply chain management where Internet of Thing term was coined for the very first time in 1991 [1][7]. IoT can provide supply chain system with real time insight of every process and transaction. The use of smart sensors and RFIDs will not only enable effective tracking of shipments as well as it will make it easy to control and manage mobile assets. It would also help in generating more business opportunities by producing analytical results on gathered information to sell goods based on this specific information.

It seems that these applications are just beginning of a big industry in computing. Moreover, this rapid development in applications shows that in the near future there will be a stable and steady stream of innovative applications and services in Internet of Things [2][3][25].

Internet of Things calls for to think beyond traditional computing. It demands small, smart and compact devices that could replace traditional computing capabilities. RFIDs, Wireless Sensor Networks, smart readers, mobile phones, laptops and portable devices are the major technologies that would work as basic computing units for such global network. RFIDs are one of the key players in IoT enabling technologies [17][25]. RFID brings into play microchips attached to any desired object for automatic identification, tracking and wireless information transmission [1]. RFIDs are used in applications of the supply chain, retail and ports for monitoring.

TABLE I. COMPARISON OF IOTS

IoT	Computational Power	Communication Range	Data rate	Storage capacity	Communication	Battery Life	Data Security
Ethernet: **LAN IEEE 802.3** **-cross over cable**	100 baseT1	100 meters	100 Mbits/s	N/A	LAN/WAN	N/A	High
Laptops: **-Dell Inspiration i7559** **-Lenovo G70 core i7**	2.6GHz 300000 D MIPS@3.0GHz	150 m	300000 D MIPS	8GB 8.1 64 bits	Wifi Bluetooth	4-8 -4-9 hrs	High
Wearables: **-Samsung Gear s3**	1Ghz	100 m	30 to 45 mbps	4GB	4G LTE	380 mAh Li-ion	Average
Smartphones: **-Infinite Note 3 pro** **-samsung galaxy J2**	1.3Ghz 1.3Ghz	130 m	upto 50 mbps	16GB 1GB	4G LTE,bluetooth,wifi 3G,bluetooth,wifi	4500 mAh 2000 mAh	Average
Cameras: **-Sony DSLR-A900** **-Canon EOS 6D**	5.0 fps 4.5 fps	1.524m/s 1.3716m/s	4 to 640 kbps	External	wired built in wifi,gps	880 shots	Low
RFIDs: **-NFC card** **-Tags**	13.56 MHz	15m	106 to 424 kbit/s	Upto 8 kb	Wireless	N/A	Low
WSN: **-open wireless sensor**	90 mips	upto 750 fet	upto250 kbps	Application dependent	Wireless Wifi	Application dependent	Low
Zigbee: **Home automation**	z-wave 90mips	upto 750 feet	upto250 kbps	Application dependent	Wireless IEEE 802.15	Application dependent	Low

Moreover, Wireless Sensor Network has turned out to be another pivot enabler for IoT. WSN uses small and intelligent computing nodes for creating a network for sensing and transmitting information from a given application field to end user's destination [25][45][51][52]. In order to monitor and accomplish real-time data, thousands of nodes are deployed for a specific set of applications. WSNs offer solutions to a wide range of applications such as industrial power control, environmental monitoring, medical Instrumentation and homeland security. Together with RFIDs, WSN is expected to get hold of highest share in key enabling technologies of IoT.

Another widely used class of technology in IoT vision 2020 is wearable computing devices that would take the personal computing to new directions. It is expected that in everyday use, wearable computing devices will frequently be used in the areas of health, education, reservation, sports, entertainment, management and controlling of resources.

Use of these devices in healthcare applications where these devices are utilised to monitor blood pressure, heart rate, and predict different diseases by using computer vision and artificial intelligence [28][29][47]. In connection with all these above discussed IoT technologies, Table 1 shows a comparison of different types of IoT devices based on their attributes such as computational power, communication range, data rate, storage, battery life and data security. The table also demonstrates that IoT devices hold the highest degree of heterogeneity and this heterogeneity is not only in device hardware, but also in their data rates, types of data generated and communication capabilities. Although, there are numerous questions that visionaries and researchers have to work out for making such applications more efficient and reliable.

Today, all these technologies work very efficiently for a specific set of applications, but they neither collaborate, nor share resources for distributed problem solving. However, in the sense of enabling technologies, there are multiple challenging areas such as device identification, interaction mechanism, standardization issues and inter devices collaboration for these heterogeneous devices [3][11][17][21][36][52].

The vision of Internet of Things seems to let small devices generate consistent data. This data will then help the decision makers to take a decision based on the enormous amount of data collected from different heterogeneous devices over a period of time [2][43]. Nevertheless, this also means IoT enabled devices will be producing data at a very high rate, which would need a huge amount of storage space. Other than the data there are multiple other challenges posed by these devices. IoT devices pose highest levels of heterogeneity problems with respect to device nature, manufacturers, communication standards, and deployed application. Second, the data generated is of multiple types and semantically different contexts. Processing and managing such data of different contexts in order to solve a set of problems for IoT applications is another leading challenge. Third, devices in the internet of things will be utilizing multiple encoding decoding mechanisms. Therefore, it will be challenging task for data management process to handle this change in encoding decoding methods. Fourth, archiving such huge amount of data for future use of IoT applications is also a foremost challenge to address primarily.

The heterogeneous nature of IoT devices sets various added challenges for data management, such as data abstraction, classification, compression, access control, archiving, interoperability, privacy and protection [10][45][54]. The ensuing need is for mature data acquisition and processing systems. Further, we need efficient data management frameworks for semantic-based data extraction from IoT devices and processing them accordingly. It is also important to note that no mature data management solutions to address above mentioned IoT centric challenges exist today. Even though, data management techniques for individual computing paradigms are performing well. But, we need to integrate them to formulate solutions for the data management requirements of Internet of Things network [8][12][14][20]. The rest of the paper is organised as follows:

Section II explores key challenges in IoTs. Section III discusses related work. Section IV presents proposed data management framework, while Section V articulates the conclusions.

II. Key Challenges

Rapid growth in IoT applications also gives birth to issues and challenges that still need to be addressed. A lot of work has already been done in this regard, but still needs sufficient research to mitigate challenges faced. In this connection, following are the key challenges confronted in IoT data management [3][4][5][11][21][36].

Figure 1 shows a diagrammatical representation of identified challenges. The figure gives a brief overview of the services provided by IoT and the data management related issues present. The outer layer represents different IoT services such as smart home, healthcare, industrial automation and city traffic management. On the other hand, inner layer represents current challenges identified in IoT services. These challenges mainly involve data integrity, data heterogeneity, knowledge management and data analysis tools. The detail of the challenges is given as follows:

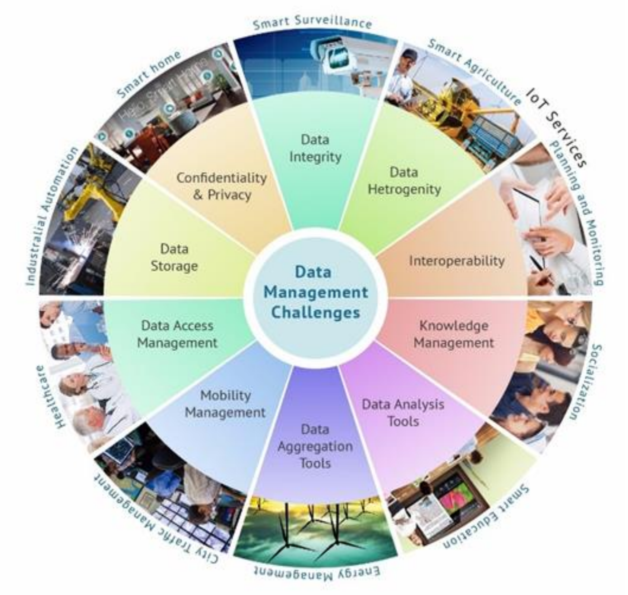

Fig. 1. IoT Services and Data Management Challenges

A. *Standardization:* Industrial IoT applications still lack the global standards that IoT enabled devices needs to follow. These standards are very crucial and will play a fundamental role for interoperability and scalability of IoT on a global scale [5][7][11][13]. Researchers, practitioners and organizations are still working to set standards for IoT. Key organizations working for setting standards include IEEE, ANSI, European Committee for Electro-Technical Standardization, and China Electronics Standardization Institute. After setting globally recognised standards industries can implement industrial applications reliably and successfully [7]. Moreover, these standards will also make it easy to convince industrialists to use IoT enabled technologies. However, this is not an easy task to standardise billions of heterogeneous devices being manufactured in different parts of the world. These IoT enabled devices shall use standard protocols and encryption techniques in order to make interoperability possible.

B. *Data storage and management:* One of the major research concerns for the next few years could be how to store data produced by objects more than the human population. In order to cater to this challenge in IoT applications, we need to employ mechanisms and frameworks to gather, store and manage data generated in IoT processes. In addition to this, we need analysis tools which may help analyse the produced data for better industrial decisions and enhancing the performance and production of different applications [6][16][35].

C. *Confidentiality and privacy:* As IoT works on sensing, tracking and connecting everyday life objects used by humans, this adds more concerns regarding privacy and information leakage [3][4][8][10][21][22]. This also produces a large amount of personal user information and hence creates the need for providing confidentiality and privacy. This requires following secure mechanisms for data collection and data access. Mechanisms should also employ that when and at what extent of data should be collected.

D. *Integrity:* One of the significant issues in any data centric environment is data integrity [26][27][29][30]. Sensing devices must gather and share only data essential to perform a required operation and assure that data is not kept or shared indefinitely. Data collection and sharing mechanisms must employ scale of integrity meaningfully with some standard procedures and rules. Data integrity is an important factor in almost any data and computation related context with the proliferation.

E. *Energy constraints:* For smooth and nonstop IoT operations, devices will need an uninterrupted power supply. These devices are not rich enough in terms of memory, processing power and energy. So, these energy constrained devices must be deployed with light weight mechanisms for device discovery, communication and invocation [39][40][41].

F. *Device mobility and heterogeneity:* Mobility of smart devices is one of the key factors in the rise of IoT. But managing this tremendous amount of mobile devices becomes an imperative challenge as well [42][43][45][46]. Internet of Things employs the use of these devices with a higher rate of mobility and heterogeneity, so it must utilise systems that support these device attributes.

G. *Device security and backup:* Mobile devices of IoT infrastructure must be secured against attacks because these nodes may be easiest victims of the attack and can effortlessly provide a gateway to an adversary to get into the system for malicious activity. This provides an attacker with the facility to disrupt whole IoT operations considerably [51][52].

H. *Availability:* Availability of IoT services must be ensured due to their critical application nature. Unavailability of these services will not only decrease overall performance, but it can also provide the attackers with the facility to launch different types of attacks against critical applications such as smart city, smart home and smart industries [27].

I. *Internal adversaries:* The significance of internal IoT adversary attack is superior to the external attacker because the internal adversary is part of IoT services and has good knowledge of different IoT components. It is relatively easy for an internal adversary to compromise some system parts or physically damage devices to disrupt services and in the long run, this can threaten the whole operation. Independent multi-layer security mechanisms should be utilised so that if the adversary is able to compromise some part of the infrastructure, then it should not affect the rest of the security methods.

In contrast with all disscussed challenges, Figure 2 presents the different IoT data management challenges, which are represented horizontally whereas; vertical lines represents the number of data management models found in the literature. While going through the literature it was observed that there was less research on most of the data management challenges such as data aggregation, data analysis and data storage. On the other hand, there was even less research on areas such as data privacy, knowledge creation, context management and data heterogeneity. Figure 2 shows this relation of data management models and their work towards different challenges.

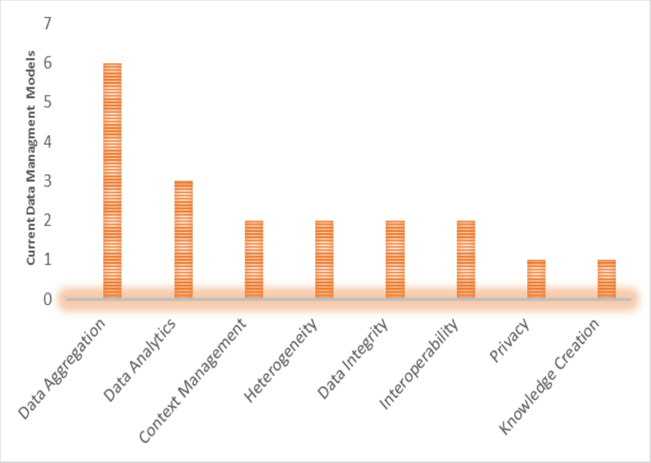

Fig. 2. Current Contribution Towards Identified IoT Data Management Challenges

III. Related Work

The core reason for the huge amount of data generated by Internet of Things enabled devices is the increasing number of internet-enabled devices used for different purposes by individuals, businesses and governments. These devices are used for the purpose of data analysis, information management, knowledge creation and knowledge management. This helps in effective policy and decision-making. Consequently, this large amount of data engendered by the IoT devices requires more and more computational power to process. Further, data generated by IoT devices in different application domains are time critical. Therefore, processing such data in a timely manner is very demanding on Internet of Things, But at the same time, considering device capabilities and context is also equally important for Complex Event Processing (CEP). In this section, we briefly discuss the most important research outcomes for IoT data management as follow:

TABLE II. COMPARISON OF CURRENT IOT DATA MANAGEMENT FRAMEWORKS

Framework/ Model	Data Aggregation	Data Analytics	Context Management	Heterogeneity	Data Integrity	Interoperability	Privacy	Knowledge Creation
COIB-Framework [2]	✓	✓	✘	✘	✘	✘	✘	✓
Service-oriented data management framework [6]	✓	✘	✘	✓	✘	✓	✘	✘
A Policy-Based Coordination Architecture [19]	✘	✘	✓	✘	✘	✓	✘	✘
A Data-Centric Framework [20]	✓	✘	✓	✓	✘	✘	✘	✘
A Large-Scale Object-Based Active Storage [33]	✓	✓	✘	✘	✓	✘	✘	✘
An Intelligent Storage Management System [34]	✓	✓	✘	✘	✘	✘	✘	✘
An architecture based on Internet of Things to support mobility [43]	✓	✘	✘	✘	✓	✘	✓	✘

Farzad et al. in [20], propose a framework to develop and deploy IoT applications in the cloud. The designed framework benefits from the current modules of Aneka and additionally pays attention towards novel features needed for IoT applications. For communication between data sources and Aneka platform, a lightweight protocol MQTT is utilised. The proposed framework has three major elements, i.e. application manager, cloud manager and data source manager. The application manager is partitioned into components like, Application Composer, Application Monitor, Scheduler and Load Balancer. These components provide the user with application side functionalities such as creation, scheduling and monitoring of applications.

Further, cloud manager handles aspects related to cloud storage. This component performs duties like allocating cloud resources, scaling resource as per need and monitoring the distributed resources to monitor resources and overall performance. Moreover, the data source manager is bridging component between framework and data sources. In order to deal structured and unstructured data, the framework uses structured and unstructured data source manager separately. The components of data source manager are able to filter specific data sent from sources to be delivered to end user’s application. For network delay reduction, these data source managers must be utilised in close proximity to the data resources. The authors also deployed a test bed with five virtual machines on Amazon AWS for performance assessment of the proposed framework.

Internet of things is an infrastructure, which will be fed by numerous heterogeneous devices in the form of data. This data needs to be in a format compatible with the storage system. However, data generated by IoT enabled devices has redundancy, anomalies and different level of abstractions. Consequently, the data generated is structured, semi-structured and unstructured. As a solution to these problems, Mishra et al. in [02] propose Cognitive Oriented IoT Big-data Framework (COIB-framework).

The proposed framework encompasses different components such as physical devices, logical IoT segments, IoT big data aggregators, IoT big data classifiers, HBase storage, IoT big data analysis and cognitive decisions. Initially, raw data is produced from physical devices which act as a data source for the whole operation. Because data at this stage is redundant, inconsistent and anomalised, therefore IoT big data aggregators are utilised to perform data fusion on this data. This step removes inconsistencies and anomalies from data to produce standard data semantics. Then, IoT big data classifiers generate clusters from this data based on their different attributes. Afterwards, classified data is stored by using HBase storage system. Now, the data can be analysed by employing cognitive and computational intelligence (CI) tools. This stage is called IoT big data analysis. As a result of the whole process, effective decisions and plans are formulated for different application sets. Authors [02] have also mentioned the use of data centres for large scale application implementation of their model where data centres will perform operations of aggregation, classification and storage operations on collected raw data.

In order to meet increasing needs of the urban population, the idea of IoT is very swiftly shifting the paradigms of urban human life. This requires making cities smart enough to enable all the operations such as education, traffic, energy management and health care can be smartly managed. This will result in having easy and timely access to real time information. This information will help in taking critical decisions necessary to provide better services to people. Gubbi et al. in [18] proposed a noise mapping architecture for both fixed (Wireless Sensor Network) and mobile infrastructures (smart phones, vehicles with smart devices/sensors and other handheld devices) in smart cities. The proposed architecture consists of the three tiers i.e. bottom tiers (consisting of sensor nodes mounted on street lights, buildings, traffic signals, etc.), middle tier (made up of relay nodes capable of collecting, buffering and transmitting information received from bottom tiers towards next tier) and top tier (acting as a gateway for sending information received from the middle tier to the cloud). Authors have utilised low-density data mode and high-density modes as the two modes for network architecture. Furthermore, cloud-computing platforms such as Microsoft Azure and Manjrasoft Aneka are utilised for interaction and real time analytics on data from IoT enabled devices. Data collected from fixed or mobile infrastructure is stored on cloud storage along with timestamps for received data. The paper also presents a noise mapping case study for the progression of city services.

To solve problems such as latency, remote policy updates, mobility and global system view, Jorge in [19] proposed a Distributed Complex Event Processing (CEP) architecture to process data from different devices bearing in mind the type and location of the sending device. To solve latency problem, data should be processed near the device or in the device. To serve this purpose, the authors defined the rules and the coordination policies, which employ that where and at what time the data is to be analysed. The proposed architecture is named as GiTo. In this architecture to make timely decisions, device attributes (such as location, battery life and location) are also well thought-out and Distributed CEP engine keeps an eye on policies and critical events observed on devices.

GiTo engine architecture has eight major components. These architectural components are Context Manager (responsible for maintaining device current context), CEP Engine, Connection Manager (manages connections between devices), Handover Manager (For keeping network connection state and active communication), Registry Manager (preserves cluster information for device), Database Manager (for exploiting system knowledge base) and HAL (for resolving platform compatibility issues in devices).

Roman et al. in [21], are more focused towards activities of data association, inference and knowledge discovery process in IoT big data management. Authors also provide précised future directions for IoT knowledge discovery.

In one more work, a cognitive IoT framework has been presented to enhance the capabilities of semantic derivations from collected data, knowledge management, discovery, and decision making process [22].

In connection with the existing works done in IoT data management area, Table 2 illustrates current IoT data management frameworks along with their contribution towards different challenges present in IoT. Further, the table also shows which challenges still need more considerable attention. It is also clear that no current data management framework reflects a conceptual solution to all the identified challenges unaccompanied.

IV. Proposed Data Management Framework

In this section, our proposed data management framework has been discussed. Data management activity is divided into multiple stages. Breaking down the data management activity into different layers leads to easiness, completeness and scalable functionality. The proposed framework contributes with wider context towards collection, management and analysis requirements of the internet of things. The proposed framework is organised into nine layers. The framework incorporates layers such as data collection layer, fog computing layer, integrity management layer, security layer, data aggregation layer, data analysis layer, data storage layer, application layer and archiving layer. Every layer of framework stack contributes for next layer of data management process. Proposed framework layers are explained in the followings section (Figure 3):

1) Data Collection Layer: the First layer in proposed framework is data collection layer. This layer works as a pass-through layer that gathers data coming from different sources and directs it to upper layers for processing [7][15][18]. Data collection layer primarily deals with numerous heterogeneous devices which are used to sense and generated data in different environments. Major devices involved in this layer can be sensors, smart devices, RFIDs, wearables, barcode readers and surveillance devices. These devices act as distinct data feeds for data collection layer. Further, this collected data can be in different forms and formats. Depending on the application nature, data collection can be centralised or distributed. This layer carries data for next layer up i.e. Fog computing layer.

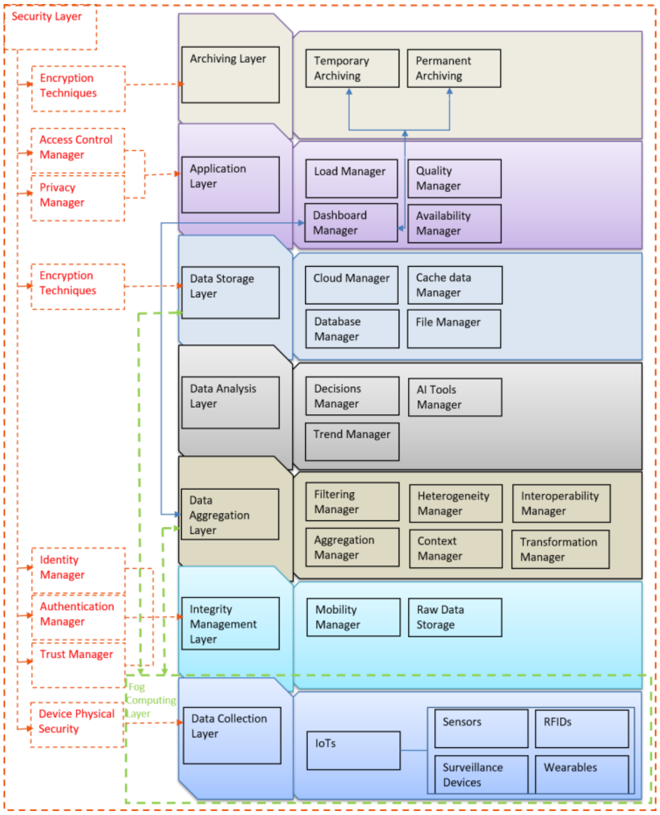

Fig. 3. Data Management Framework

2) Fog Computing Layer: Futuristic and time critical applications essentially demand the analysis of data be performed nearby its point of generation rather sending the data to cloud every time for analysis and decision making [55, 56, 57, 58]. Consequently, this calls for shifting data management functionalities closer to data generating devices. By keeping this need of data management in mind, the fog-computing layer has been anticipated in the proposed model. This layer provides devices with facilities to process, analyse and partially store data on / nearby edge nodes. In our proposed model, fog computing layer is mainly associated with data collection, data aggregation and data storage layer. However, in order to accommodate data management functionalities at edge nodes, devices must have upright and dominant computational power, memory, battery life and all other required resources. Moreover, only time critical data is aggregated and analysed on devices otherwise it is sent to the cloud for the long-term analysis and storage.

3) Integrity Management Layer: This layer is responsible for the integrity of whole data management process [27]. Major components of this layer are raw data storage and mobility manger.

Moreover, for addressing gaining concerns of device mobility in IoT, mobility manager is mostly attentive towards features associated with device mobility. Mobility manager also takes into account the impact of device mobility on the context of data sent by the device. In addition, mobility manger must employ schemes to support service mobility, session mobility and personal mobility. This module should also manage data handoff between devices. Raw data storage module manages storage of raw, enormous and continuous stream of data prior to any data management operation performed on this received data. This module must be equipped with techniques and tools to store huge influx of data by maintaining data repositories. This also involves maintaining metadata and indexes for the stored data.

In order to take benefits of data management process integrity management layer should essentially take care of issues concerning authenticity, integrity and availability. This will not only diminish data management overhead but also influence the quality of analysis and decisions taken based on this provided data.

4) Data Aggregation Layer: One of the core intentions of this layer is data size reduction for improved storage, organization and transmission of data [2]. For this reason, data aggregation layer is concentrated towards real time summarization and merging of the data. The key modules of this layer are filtering, heterogeneity, interoperability, aggregation, context, and transformation manager.

Data received from integrity management layer is raw, redundant and very huge. It must be pre-processed before offered to advance layers of the data management process for further processing. Filtering is a fundamental stage of data reconstruction and event processing. Filtering helps to make raw data relatively more meaningful and reducing noise from data. Filtering manager sets filtering conditions for received data and pre-processes data. These filters may be temporary, permanent or based on the frequency of necessity. Further, these filters can be set according to user requirements and choice for the applications.

Another component of data aggregation layer is heterogeneity manager. Heterogeneity of IoT devices lays new challenges to their management due to the absence of unifying approaches. In this regard, heterogeneity manager is responsible for handling device heterogeneity, data heterogeneity and semantic heterogeneity. Simultaneously, this heterogeneity also forces shift for the necessity of transforming data coming in various types and varied sampling intervals into single or multiple predefined data types for efficiency and ease of further processing. Transformation manager is also responsible for transparently transforming data received into the user’s application view format. The leading activities involved at this stage are data splitting, merging and sorting.

While IoT connects various heterogeneous devices, their interoperability is very crucial for seamless communication between devices and services. In this regard, standards for device representation, searching and access must be defined. This also needs one/multiple common communication languages for information exchange between devices from different vendors. Interoperability manger has to handle technical interoperability, semantic interoperability, syntactic interoperability and cross-domain interoperability.

Furthermore, IoT demands the context–aware data gathering and management. In this connection, context manger looks for the most suited devices that could generate most relevant data for user application. Context manger maintains context information for device or group of devices gathering real time data for some specific context. In this way, context specification support of data can help developing new value added services for users. Consequently, contextual data can be delivered to the user very easily and effectively. This component also works for context representation and modelling. Context manager should also be able to identify new contexts drawn from previous data.

After performing above specified processes in data aggregation layer, aggregation manager further processes and aggregates data into the form appropriate for further analytics. Aiming at this purpose, a set of data aggregation tools for data validation, processing and aggregation in the specified format are utilised. Aggregation manager also chooses data that encounter specific principles and standards.

5) Security Layer: Retaining automated security in IoT still remains in the spotlight due to the significance of security needs. In our proposed framework, security layer guarantees security provision to all layers involved in data management process. In this way, the security layer is linked with all layers in the model and intended to meet security requirements at each individual layer. According to the functionality of different layers in the model, this layer provides respective security tools for securing that layer's operations. Main components of security layer are trust manager, authentication manager, identity manger, device physical security, encryption techniques, privacy manager and access control manager [50][53].

Integrity management layer is supported with security modules such as identity, authentication and trust manager. Validating data sources is one of the crucial tasks for IoT data management. In this regard, Identity manger will be responsible for the identification of the data sources. Every device involved is assigned an ID and Identity manager treats data conferring to its identity. Moreover, this identity information is interrelated with the authentication process incorporated by authentication manager whereas trust manager maintains trust level for all devices engendering data. This trust level of devices is based on their data correctness, completeness and timeliness. Trust manager periodically re-computes and updates trust levels of devices by taking into account their current data activity.

At the application layer, access control manager and privacy manger are employed for fulfilling security requirements at this layer. Privacy manger is mostly directed towards defining application privacy policies [50][53]. It also provides protection mechanisms to end users form exposure to privacy risks whereas access control manager works for controlled data access of devices and users. The main responsibilities involve data ownership, secure data sharing, distributed data access and data access permissions.

In order to provide security at data storage layer and archiving layer, various encryption techniques and tools are exploited. From the perspective of time, encryption mechanisms used in archiving layer can be computationally expensive because of non-time sensitivity. Whereas, in order to meet run time data retrieval and storage in IoT applications, encryption techniques at data storage layer should not be computationally expensive but at the same time these techniques should ensure data security at this layer.

6) Data Analysis Layer: This layer augments significance to the data gathered by analysing it to engender smart decisions and analysis [2][6]. This layer will also fulfill user requirements regarding on demand user analysis and run extraction tools for desired information. This would provide users with actionable information according to the situation for making timely effective decisions and collaborate with other users and applications. Further, this layer must provision analytical data support for all types of IoT environments such as off line, dynamic and real time environments.

Data analysis layer is divided into three modules, i.e. decision manager, trend manager and AI manager. The first module of data analysis layer is decision manager who is responsible for driving decisions from the current data received from lower layers. If a user has to make some decisions related to a problem area, this module will consider current and historical data relevant to this problem and generate some decisions which seem to be most appropriate for the problem. Furthermore, decision manager will periodically make some strategic decisions based on acquired data and share these decisions timely to respective organizations and users. This module will also comfort users in contextual decision making, resulting in quicker attainment of organizational objects. Moreover, decision manger can be helpful in applications such as stock exchange, agriculture, weather forecasting and real state whereas trend manager helps understanding current trends and user interests in different application areas. Trend manager can also find latest national and international trends related to youth, politics, sports and social interest. This also benefits finding data trends for the user that which type of data is more utilised by the user. This will facilitate organizations to understand user requirements and current interest better and create products according to this trend analysis. Additionally, this will improve market profit and competitive intelligence.

The artificial intelligence manger in data analysis layer plays a very fundamental role. AI manager employs machine learning, neural network and deep learning tools to analyse data.AI manger will providree IoT with the continuous learning of new analytical models and algorithms on available data. Further, AI manager works for creating automated intelligent systems to fulfill analysis and management requirements of IoT data. However, the correctness and speed of AI manager must be improved for IoT to perform according to real time IoT applications' needs.

7) Data Storage Layer: Due to continuous generation of huge amount of data in different varieties and quantities, the necessity for standardised and efficient mechanisms for data storage is more imperative than ever. Data storage layer is responsible for real time data storage as data is produced [6][33]. This layer also resolves data storage location problems by taking into account nature of data and application requirements. Another aspect that needs consideration at this layer is data storage format for different types of data provided by lower layers. Furthermore, this layer also upholds indexing, catalogues and semantic metadata of stored data for timely retrieval. The major components of this layer are a cloud, cache, database, and file manager.

Cloud manager will look after data storage aspects regarding cloud storage. This storage can be used mostly by the organizations wishing to use cloud storage as a service besides managing their own storage infrastructure. This will provide flexibility and scalability of data storage to such organizations. For timely and fast provisioning of contents to IoT applications, the cache manager is directed towards cache organization and maintenance. For dissimilar types of data, the cache manager will be responsible for defining policies for caching heterogeneous data. These policies may be general, time-based and location-based. Cache manager will also classify cached data into a number of categories according to user's application requirements and maintain this information.

8) Application Layer: Application layer will be focused towards providing services to end users and governs data flow. Application layer also performs the duty of load balancing. Further, this layer is responsible for maintaining the quality of service in terms of data for the end user [6][9][15][17]. This layer also looks at the availability of data for application domains. The main modules of this layer are load, quality, dashboard, access control, and availability manager.

The support, manageability and continuous supply of high data traffic demand for load balancing mean to acquire data from sources. Load balancing manager plays very important role in scalability, reliability and enhanced performance of IoT data management lifecycle. This module will employ routing policies and algorithms for distributing data requests at sources such that data acquiring load is distributed across available sources. Furthermore, after a fixed or arbitrary interval this component will search for over-utilised and under-utilised resources for effective load distribution. Consequently, load balancing increases device life time and availability for energy constrained devices in IoT.

In order to accommodate consistent and continuous data generation process, availability of IoT devices is very imperative. Data availability makes possible smooth, timely and uninterrupted data management lifecycle. Availability manager looks for the availability of the devices and if some of the data sensing devices for a particular application are down, availability manger explores some other devices which can send data instead of unavailable sources. Availability manager not only enlists availability of devices but it also tries to increase their availability by taking into account their resources. These two modules of application layer can coordinate such that availability manager keeps track of available devices and their resources in terms of available computing power, memory and energy. Then, this information is shared with load balancing manager to distribute data load. Subsequently, this coordination results in reduced service delay, minimum down time and long term availability of data sources.

The quality of data is also critical for social and commercial impacts on different application domains of IoT. In this regard, IoT data should preserve properties such as completeness, correctness and quality of information. Quality manager practices tools and techniques to encounter data quality for applications. The quality manager selects quality metrics. Further, testing of devices, platforms and corresponding technologies is also performed by this module. In addition, dashboard manager helps users to manage their application dashboards to interact, monitor and visualise their preferred contents and services. It also facilitates users by providing real time custom dashboards of user's choice. In proposed framework, the dashboard may coordinate with data aggregation layer for on demand data aggregation for users.

9) Archiving Layer: Another important aspect of IoT data management is to archive such huge amount of data generated by the devices. Archiving layer will be responsible for managing growing archiving needs of IoT data with scalable infrastructure. This layer maintains indexes for effective and timely data search. This layer will employ mechanisms so that data is not overwritten and altered. Archiving layer is further divided into two modules, i.e. temporary archiving and permanent archiving.

Proposed framework caches most frequently accessed IoT data in data storage layer. However, data with relatively less access frequency will be archived provisionally by the temporary archiving module. This module will manage data for short term data retention. Further, this module will define and manage policies to select low priority and aging data from temporary archives so that it can be sent to the permanent archiving module. This also involves making decisions concerning preservation requirements for various types of available data. Whereas, the permanent archiving module is focused towards preserving the data for the indefinitely long time that is occasionally requested and remains unused in the day to day operations. This module employs redundancy and cryptographic techniques for security, durability and cost effectiveness. Furthermore, this module also controls access to these archived contents.

V. Conclusions and Future Directions

Growing technological evolution of computing devices, IoT has become a vital part of modern computing world especially for the large-scale computing infrastructures. Internet of Things has many applications in different areas. However, current solutions for the data management in IoTs addresses only the partial aspects of the cloud centric IoT environment with special focus on sensor networks, which is only a subset of the global IoT space. Although, mobile devices such as smart phones, surveillance devices and other smart handheld/wearable devices are generating data at much higher rate, but their data management concerns are still a point of concern. Solutions to manage and utilise the massive amount of data that is being generated by these objects are yet to mature. Industry wide global standards, unified communication protocols, highly enhanced security aspects and middleware problems are left for future work.

References

[1] Gubbi, Jayavardhana, et al. "Internet of Things (IoT): A vision, architectural elements, and future directions." Future Generation Computer Systems 29.7 (2013): 1645-1660.

[2] Mishra, Nilamadhab, Chung-Chih Lin, and Hsien-Tsung Chang. "A cognitive adopted framework for IoT big-data management and knowledge discovery prospective." International Journal of Distributed Sensor Networks 2015 (2015): 6.

[3] Weber, Rolf H. "Internet of Things–New security and privacy challenges." Computer Law & Security Review 26.1 (2010): 23-30.

[4] Elmaghraby, Adel S., and Michael M. Losavio. "Cyber security challenges in Smart Cities: Safety, security and privacy." Journal of advanced research 5.4 (2014): 491-497.

[5] Miorandi, Daniele, et al. "Internet of things: Vision, applications and research challenges." Ad Hoc Networks 10.7 (2012): 1497-1516.

[6] Fan, Tongrang, and Yanzhao Chen. "A scheme of data management in the Internet of Things." *2010 2nd IEEE InternationalConference on Network Infrastructure and Digital Content*. IEEE, 2010.

[7] Da Xu, Li, Wu He, and Shancang Li. "Internet of things in industries: A survey." *IEEE Transactions on Industrial Informatics* 10.4 (2014): 2233-2243.

[8] Kumar, J. Sathish, and Dhiren R. Patel. "A survey on internet of things: Security and privacy issues." International Journal of Computer Applications 90.11 (2014).

[9] Tan, Lu, and Neng Wang. "Future internet: The internet of things." *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*. Vol. 5. IEEE, 2010.

[10] Henze, Martin, et al. "User-driven privacy enforcement for cloud-based services in the internet of things." Future Internet of Things and Cloud (FiCloud), 2014 International Conference on. IEEE, 2014.

[11] Bandyopadhyay, Debasis, and Jaydip Sen. "Internet of things: Applications and challenges in technology and standardization." *Wireless Personal Communications* 58.1 (2011): 49-69.

[12] Kitchin, Rob. "The real-time city? Big data and smart urbanism." GeoJournal 79.1 (2014): 1-14.

[13] Atzori, Luigi, Antonio Iera, and Giacomo Morabito. "The internet of things: A survey." *Computer networks* 54.15 (2010): 2787-2805.

[14] Botta, Alessio, et al. "Integration of cloud computing and internet of things: a survey." *Future Generation Computer Systems* 56 (2016): 684-700.

[15] Gaikwad, Pranay P., Jyotsna P. Gabhane, and Snehal S. Golait. "A survey based on smart homes system using Internet-of-things." *Computation of Power, Energy Information and Commuincation (ICCPEIC), 2015 International Conference on*. IEEE, 2015.

[16] Bohli, Jens-Matthias, et al. "SMARTIE project: Secure IoT data management for smart cities." *Recent Advances in Internet of Things (RIoT), 2015 International Conference on*. IEEE, 2015.

[17] Al-Fuqaha, Ala, et al. "Internet of things: A survey on enabling technologies, protocols, and applications." *IEEE Communications Surveys & Tutorials* 17.4 (2015): 2347-2376.

[18] Jin, Jiong, et al. "An information framework for creating a smart city through internet of things." *IEEE Internet of Things Journal* 1.2 (2014): 112-121.

[19] Fonseca, Jorge, Carlos Ferraz, and Kiev Gama. "A policy-based coordination architecture for distributed complex event processing in the internet of things: doctoral symposium." *Proceedings of the 10th ACM International Conference on Distributed and Event-based Systems*. ACM, 2016.

[20] Khodadadi, Farzad, Rodrigo N. Calheiros, and Rajkumar Buyya. "A data-centric framework for development and deployment of internet of things applications in clouds." *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on*. IEEE, 2015.

[21] Roman, Rodrigo, Jianying Zhou, and Javier Lopez. "On the features and challenges of security and privacy in distributed internet of things." Computer Networks 57.10 (2013): 2266-2279.

[22] Sicari, Sabrina, et al. "Security, privacy and trust in Internet of Things: The road ahead." Computer Networks 76 (2015): 146-164.

[23] Vasilomanolakis, Emmanouil, et al. "On the Security and Privacy of Internet of Things Architectures and Systems." 2015 International Workshop on Secure Internet of Things (SIoT). IEEE, 2015.

[24] Abomhara, Mohamed, and Geir M. Køien. "Security and privacy in the Internet of Things: Current status and open issues." Privacy and Security in Mobile Systems (PRISMS), 2014 International Conference on. IEEE, 2014.

[25] Chen, Chien-Ming, et al. "RCDA: recoverable concealed data aggregation for data integrity in wireless sensor networks." IEEE Transactions on parallel and distributed systems 23.4 (2012): 727-734.

[26] Luo, Wenjun, and Guojing Bai. "Ensuring the data integrity in cloud data storage." 2011 IEEE International Conference on Cloud Computing and Intelligence Systems. IEEE, 2011.

[27] Bowers, Kevin D., Ari Juels, and Alina Oprea. "HAIL: a high-availability and integrity layer for cloud storage." Proceedings of the 16th ACM conference on Computer and communications security. ACM, 2009.

[28] Di Pietro, Roberto, and Luigi V. Mancini. "Security and privacy issues of handheld and wearable wireless devices." Communications of the ACM 46.9 (2003): 74-79.

[29] Falk, Jennica, and Staffan Björk. "Privacy and information integrity in wearable computing and ubiquitous computing." CHI'00 extended abstracts on Human factors in computing systems. ACM, 2000.

[30] [30] Frikken, Keith B., and Joseph A. Dougherty IV. "An efficient integrity-preserving scheme for hierarchical sensor aggregation." Proceedings of the first ACM conference on Wireless network security. ACM, 2008.

[31] Chen, Fei, and Alex X. Liu. "Privacy-and integrity-preserving range queries in sensor networks." IEEE/ACM Transactions on Networking 20.6 (2012): 1774-1787.

[32] Yang, Jiachen, et al. "Multimedia cloud transmission and storage system based on internet of things." Multimedia Tools and Applications (2015): 1-16.

[33] Xu, Quanqing, et al. "A large-scale object-based active storage platform for data analytics in the internet of things." Advanced Multimedia and Ubiquitous Engineering. Springer Berlin Heidelberg, 2016. 405-413.

[34] Kang, Jun, Siqing Yin, and Wenjun Meng. "An Intelligent Storage Management System Based on Cloud Computing and Internet of Things." Proceedings of International Conference on Computer Science and Information Technology. Springer India, 2014.

[35] Fan, Tongrang, and Yanzhao Chen. "A scheme of data management in the Internet of Things." 2010 2nd IEEE InternationalConference on Network Infrastructure and Digital Content. IEEE, 2010.

[36] Barnaghi, Payam, Amit Sheth, and Cory Henson. "From Data to Actionable Knowledge: Big Data Challenges in the Web of Things [Guest Editors' Introduction]." IEEE Intelligent Systems 28.6 (2013): 6-11.

[37] Jara, Antonio J., et al. "Yoapy: A data aggregation and pre-processing module for enabling continuous healthcare monitoring in the internet of things." International Workshop on Ambient Assisted Living. Springer Berlin Heidelberg, 2012.

[38] Korteweg, Peter, et al. "Data aggregation in sensor networks: Balancing communication and delay costs." International Colloquium on Structural Information and Communication Complexity. Springer Berlin Heidelberg, 2007.

[39] Looga, Vilen. "Energy-awareness in large-scale internet of things networks." Proceedings of the 2014 workshop on PhD forum. ACM, 2014.

[40] Jammes, Francois. "Internet of Things in Energy Efficiency: The Internet of Things (Ubiquity symposium)." Ubiquity 2016.February (2016): 2.

[41] Chatzigiannakis, Ioannis, Dimitrios Amaxilatis, and Spyros Livathinos. "A collective awareness platform for energy efficient smart buildings." Proceedings of the 19th Panhellenic Conference on Informatics. ACM, 2015.

[42] Zorzi, Michele, et al. "From today's intranet of things to a future internet of things: a wireless-and mobility-related view." IEEE Wireless Communications 17.6 (2010): 44-51.

[43] Valera, Antonio J. Jara, Miguel A. Zamora, and Antonio FG Skarmeta. "An architecture based on internet of things to support mobility and security in medical environments." 2010 7th IEEE Consumer Communications and Networking Conference. IEEE, 2010.

[44] Shon, Taeshik, et al. "Toward advanced mobile cloud computing for the internet of things: current issues and future direction." Mobile Networks and Applications 19.3 (2014): 404-413.

[45] Mantri, Dnyaneshwar S., Neeli Rashmi Prasad, and Ramjee Prasad. "Mobility and Heterogeneity Aware Cluster-Based Data Aggregation for Wireless Sensor Network." Wireless Personal Communications 86.2 (2016): 975-993.

[46] Hsiao, Yuan-Kai, and Yen-Wen Lin. "A Mobility Management Scheme for Internet of Things." Mobile, Ubiquitous, and Intelligent Computing. Springer Berlin Heidelberg, 2014. 569-575.

[47] Li, Fagen, Yanan Han, and Chunhua Jin. "Practical access control for sensor networks in the context of the Internet of Things." Computer Communications (2016).

[48] Mahalle, Parikshit N., et al. "A fuzzy approach to trust based access control in internet of things." Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE), 2013 3rd International Conference on. IEEE, 2013.

[49] Liu, Jing, Yang Xiao, and CL Philip Chen. "Authentication and Access Control in the Internet of Things." ICDCS Workshops. 2012.

[50] Mahalle, Parikshit N., et al. "Identity authentication and capability based access control (iacac) for the internet of things." Journal of Cyber Security and Mobility 1.4 (2013): 309-348.

[51] Becher, Alexander, Zinaida Benenson, and Maximillian Dornseif. "Tampering with motes: Real-world physical attacks on wireless sensor networks." International Conference on Security in Pervasive Computing. Springer Berlin Heidelberg, 2006.

[52] Ghosal, Amrita, and Subir Halder. "Intrusion detection in wireless sensor networks: issues, challenges and approaches." Wireless Networks and Security. Springer Berlin Heidelberg, 2013. 329-367.

[53] Babar, Sachin, et al. "Proposed embedded security framework for internet of things (iot)." Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE), 2011 2nd International Conference on. IEEE, 2011.

[54] Busold, Christoph, et al. "Smart and Secure Cross-Device Apps for the Internet of Advanced Things." International Conference on Financial Cryptography and Data Security. Springer Berlin Heidelberg, 2015.

[55] Bonomi, Flavio, et al. "Fog computing and its role in the internet of things." *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*. ACM, 2012.

[56] Dastjerdi, Amir Vahid, and Rajkumar Buyya. "Fog Computing: Helping the Internet of Things Realize Its Potential." *Computer* 49.8 (2016): 112-116.

[57] Bonomi, Flavio, et al. "Fog computing and its role in the internet of things." *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*. ACM, 20126.

[58] Qaisar, Saad, and Nida Riaz. *Fog Networking: An Enabler for Next Generation Internet of Things*. International Conference on Computational Science and Its Applications. Springer International Publishing, 2016.