EAI/Springer Innovations in Communication and Computing

Fadi Al-Turjman *Editor*

Edge Computing

From Hype to Reality

EAI
RESEARCH MEETS INNOVATION

Springer

EAI/Springer Innovations in Communication and Computing

Series editor
Imrich Chlamtac, CreateNet, Trento, Italy

Editor's Note

The impact of information technologies is creating a new world yet not fully understood. The extent and speed of economic, life style and social changes already perceived in everyday life is hard to estimate without understanding the technological driving forces behind it. This series presents contributed volumes featuring the latest research and development in the various information engineering technologies that play a key role in this process.

The range of topics, focusing primarily on communications and computing engineering include, but are not limited to, wireless networks; mobile communication; design and learning; gaming; interaction; e-health and pervasive healthcare; energy management; smart grids; internet of things; cognitive radio networks; computation; cloud computing; ubiquitous connectivity, and in mode general smart living, smart cities, Internet of Things and more. The series publishes a combination of expanded papers selected from hosted and sponsored European Alliance for Innovation (EAI) conferences that present cutting edge, global research as well as provide new perspectives on traditional related engineering fields. This content, complemented with open calls for contribution of book titles and individual chapters, together maintain Springer's and EAI's high standards of academic excellence. The audience for the books consists of researchers, industry professionals, advanced level students as well as practitioners in related fields of activity include information and communication specialists, security experts, economists, urban planners, doctors, and in general representatives in all those walks of life affected ad contributing to the information revolution.

About EAI

EAI is a grassroots member organization initiated through cooperation between businesses, public, private and government organizations to address the global challenges of Europe's future competitiveness and link the European Research community with its counterparts around the globe. EAI reaches out to hundreds of thousands of individual subscribers on all continents and collaborates with an institutional member base including Fortune 500 companies, government organizations, and educational institutions, provide a free research and innovation platform.

Through its open free membership model EAI promotes a new research and innovation culture based on collaboration, connectivity and recognition of excellence by community.

More information about this series at http://www.springer.com/series/15427

RESEARCH MEETS INNOVATION

Fadi Al-Turjman
Editor

Edge Computing

From Hype to Reality

Springer

Editor
Fadi Al-Turjman
Computer Engineering Department
Antalya Bilim University
Antalya, Turkey

ISSN 2522-8595 ISSN 2522-8609 (electronic)
EAI/Springer Innovations in Communication and Computing
ISBN 978-3-319-99060-6 ISBN 978-3-319-99061-3 (eBook)
https://doi.org/10.1007/978-3-319-99061-3

Library of Congress Control Number: 2018958621

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my great parents, my first lady and my little one.

Whose affection, pray, and continues encouragement have made this book and every beautiful thing in my life realized.

Fadi Al-Turjman

Preface

We are living in an era where the Edge of the Cloud is becoming a global platform for the computation and interaction between machines and smart objects, in real-time applications and for many critical tasks.

With the application areas such as smart cities, smart grids, smart cars, smart eHealth, smart supply chain, and smart homes in the Internet of Things (IoT), we can consider the Edge computing as a complete package of the smart networked objects. From this perspective, it is essential to understand the role of Edge computing which will provide a global backbone for the worldwide information sharing/processing in the near future.

There is no doubt that introducing such a new phenomenon can come with potential challenges especially in terms of security and quality of service (QoS). Therefore, it is essential to consider new enabling technologies such as wireless sensor networks (WSN), various radio technologies and cellular infrastructures (e.g., small cells), radio-frequency identification (RFID), and cloud services and architectures for performance optimization.

The objective of this book is to present a survey of existing techniques and architectures for Edge computing optimization and improvements with the presence of IoT systems. The main focus is on practical design aspects that can help in realizing such a paradigm in practice. The applications of Edge computing in IoT, evaluation metrics, constraints, and open issues about the addressed topic are included for discussion as well. This conceptual book, which is unique in the field, will assist researchers and professionals working in the area of the Cloud to better assess the proposed Edge paradigms that are already beginning to be a significant part of the global infrastructure.

Antalya, Turkey | Fadi Al-Turjman

Contents

1 The Era of the Personal Cloud: What Does It Mean for Cloud Providers? 1
Mohamed Abu Sharkh, Abdallah Shami, and Mohamad Kalil

2 Optimization in Edge Computing and Small-Cell Networks 17
Jitender Grover and Ram Murthy Garimella

3 A Comprehensive Survey on Architecture for Big Data Processing in Mobile Edge Computing Environments 33
Maninder Jeet Kaur

4 Taxonomy of Edge Computing: Challenges, Opportunities, and Data Reduction Methods 51
Kusumlata Jain and Smaranika Mohapatra

5 Applications of the Internet of Things with the Cloud Computing Technologies: A Review 71
U. Ram Jagannath, S. Saravanan, and S. Kanimozhi Suguna

6 Software-Defined Internet of Things to Analyze Big Data in Smart Cities 91
Sadia Din, Awais Ahmad, Anand Paul, and Gwanggil Jeon

7 Taking Cloud Computing to the Extreme Edge: A Review of Mist Computing for Smart Cities and Industry 4.0 in Africa 107
Eustace M. Dogo, Abdulazeez Femi Salami, Clinton O. Aigbavboa, and Thembinkosi Nkonyana

8 IoT and Edge Computing as a Tool for Bowel Activity Monitoring 133
Umit Deniz Ulusar, Erdinc Turk, Ahmet Sefa Oztas, Alp Erkan Savli, Guner Ogunc, and Murat Canpolat

9 Application of Cloud Computing and Internet of Things to Improve Supply Chain Processes 145
S. Kanimozhi Suguna and Suresh Nanda Kumar

10 A Conceptual Framework for Security and Privacy in Edge Computing 171
Bragadeesh S. A. and Umamakeswari Arumugam

Index 185

Chapter 1
The Era of the Personal Cloud: What Does It Mean for Cloud Providers?

Mohamed Abu Sharkh, Abdallah Shami, and Mohamad Kalil

1.1 Introduction

Depending on Cloud technologies for core business activities like collaboration, business continuity and operation has become a reality for businesses all over the world. Cloud technology adoption rates are increasing steadily. 78% of US small businesses will have fully adopted Cloud computing by 2020, more than doubling the current 37% [18]. The percentage grows to 90% when looking at large businesses (larger than 1000 employees) [35]. The US small and medium business (SMB) Cloud computing and services market will grow from 43 billion dollars in 2015 to 55 billion dollars in 2016 [18]. This trend is consistent in Europe as well. The percentages of small, medium, and large businesses adopting Cloud technologies in the UK are 46%, 63%, and 82%, respectively. In Germany, the percentages are 50%, 65%, and 86%. This has been the case for sometime. A game changer in the last few years is the dramatic increase in the personal use of Cloud services, so much so that Gartner has dubbed the next era as the era of the personal Cloud [22]. This means that the usage trend for individuals has moved from depending on the personal computer into depending heavily on personal Clouds. This is by no means a small or ignorable feat compared to enterprise businesses. It is large market chunk that demands attention from Cloud providers. This direction

M. Abu Sharkh (✉)
Digital Media Software Engineering, Ferris State University, Grand Rapids, MI, USA
e-mail: abusham1@ferris.edu

A. Shami
Department of Electrical and Computer Engineering, Western University, London, ON, Canada
e-mail: ashami2@uwo.ca

M. Kalil
IBM Analytics, IBM Canada, Markham, ON, Canada
e-mail: mkalil3@uwo.ca

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_1

covers a transfer away from the full range of personal computing usage patterns to a different set of patterns using Cloud services. For example, Gartner estimated that consumers' storage will grow from around 500 gigabytes (2011 estimate) to 3.3 terabytes by 2016. A third of these loads is already in the Cloud [22].

As for network loads, 2014 was the first year a majority of the world data workloads were on the Cloud (51% to 49% for data centers). Cisco expected the growth of this number to be 63% in 2017. Moreover, the estimate for network traffic was 69% for Cloud networking compared with 31% with regular data centers. Mobile date traffic is a slice that is affected largely as well [16]. Cloud mobile data traffic is projected to grow from 81% in 2014 to 90% in 2019.

Research firm IDC estimates that nearly one third of the worldwide enterprise application market will be SaaS-based by 2018, driving annual SaaS revenue to $50.8B, from 22.6B in 2013 [17]. According to a survey reported in [34], the average enterprise uses 1,427 distinct Cloud services with the average employee actively using 36 Cloud services at work. Gartner estimated the worldwide public Cloud services market at $209.2B and predicted this will pass the $250B barrier soon [23].

The Internet of Things offerings expand the horizons of data sources for a certain Cloud-based service, but it also jumps to a new dimension in terms of the data filtering and preprocessing required of the same service to get to the core value returning requests. The sheer volume of the connected devices along with the connection speed and user expectations for response from the server put a stress both on the network and on back end Cloud data center to perform up to the required level. This poses a challenge to Cloud service designers regarding how to optimize this process. You cannot expect every single data packet from 8.4B heterogeneous IoT devices (Gartner's estimate of the number of devices deployed around the world in 2017) to be informative or even relevant.

Personal Cloud spread contributed to the Cloud market size multiplication. The consumer's interest combined with technological advancements contributed to these leaps in consumption. Public Clouds – which are where most of the personal Cloud loads reside – are the most affected with this growth. Public Cloud vendor revenue grew from 26 billion dollars in 2012 to a projected 154 billion in 2017 [45].

Software as a service represented 67% of this as of 2015. Traditional IT infrastructure is definitely losing ground to Cloud solutions. The market share percentages moved swiftly from 72%, 17%, and 11% for traditional IT, public Clouds, and private Clouds, respectively, in 2011 to 53%, 29%, and 5% in 2017 [43].

Personal work loads are going in a clear direction: up to the Clouds.

In this work, we endeavor to investigate the potential impact the dawn of this new era will have on Cloud providers. The purpose is to offer a solid perspective into what changes are expected in the environment, technical requirements and usage trends. This, in turn, would aid in forming a vision for the directions providers need to go into their policies and algorithm planning.

In Sect. 1.2, we investigate the impacts on the area of resource allocation in a Cloud data center. This covers multiple venues including the request distribution effects, impacts on financial investment in Cloud resources, the optimality of the pricing plans, and energy efficiency considerations in light of the personal Cloud conditions. Section 1.3 offers a detailed discussion on Cloud client demand pattern prediction methods.

In Sect. 1.4, we introduce the impacts and the intersection points with another dominating technology trend, which is the Internet of Things. This section covers the challenges expected from numerous nodes representing the clients in that scenario. The cause and effect of the diversity and volume of these requests is probed into.

Section 1.5 strives to envision any tangible effects on Cloud brokering services. Finally, venues for future improvement are explored in Sect. 1.6.

1.2 Resource Allocation in the Cloud: Personal vs. Corporate Clients

1.2.1 X Requests vs. One Really Large One

For a Cloud provider, the resource allocation is in the core of the Cloud management process. Resource utilization and long-term planning can be really lucrative in terms of performance and financial rewards. This is generally easier to handle with enterprise clients. Despite Cloud computing having the major advantage of dynamic scaling and pay as you go as a benefit for clients, it is noticeable that precise prediction of the resource scale the client will demand would be rewarding for the client. For a large enterprise, the favorable scenario would be to rent an amount of Cloud resources as close as possible to the amount they practically require on a long-term contract with the ability to scale up or out just in case demand surges happen. As depicted in detail later, dynamic scaling is a business-enhancing feature, but it being the norm does not make sense cost-wise.

A look at some of the public Cloud price offerings confirms this idea. It is a practical and realistic assumption to expect enterprise clients to have the expected average load and maximum load estimates calculated and clear. However, with clients using personal Clouds or services with higher exposure to individuals using their own devices, this can get much trickier. For example, if we look at pricing patterns for VM instance rental for Amazon AWS and if we focus on infrastructure as a service since it often is the clearest example of the resource outsourcing and economies of scale, some interesting insights can be drawn [7].

Amazon's three charging models are on-demand, which, as the name suggests, basically reflects a pay-as-you-go system. Understandably, instances under this usage condition tend to be of higher price and priced per hour. Next is the reserved instances which provide the clients with guaranteed capacity for a fixed term. These are offered with (no upfront, partial upfront, and all upfront) payment options and on a 1-year or 3-year terms. Finally, Amazon offers the clients the opportunity to bid for the remaining unrented capacities in a specific Clouddata center. To buy these instances – called spot instances – clients determine their demands and put a bid in

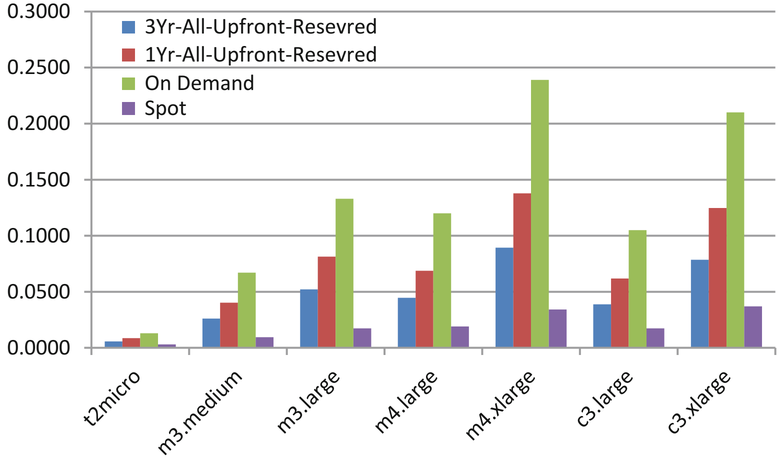


Fig. 1.1 A comparison of Cloud service prices per hour for a select set of instances offered by Amazon under different charging modes [7]

the form of the price they are willing to pay, and then the instances are sold subject to the available capacity in a data center.

Figure 1.1 shows a comparison of the prices of 4 pricing options (3-year reserved, 1 year reserved, on-demand, and spot instances based on the latest accepted bids) as of Feb 2016 for several instance types. We chose to use Linux-based instances for this comparison (since the environment change affects the price per hour). Each instance type reflects a different resource configuration. This figure along with Fig. 1.2 – which shows the percentage of the on-demand instance price for the other three pricing methods – illustrates the significance of the usage condition to the client. As for the provider side, a look at the revenue that can be generated from a specific load for a Cloud with predetermined capacity can really help us achieve and form a trend here.

The variables in play here are the prices per hour for each usage condition. α is defined to be the price per hour for 3-year reserved instances. β represents the price per hour for 1-year reserved instances. γ represents the price per hour for the on-demand instances. δ represents the price per hour for the instances sold in the spot mode based on the latest spot instance price. X represents the number of 3-year reserved instances. Y represents the number of 1-year reserved instances. Z represents the number of on-demand instances. Finally, W is the number of instances sold in spot mode in a specific hour. The hourly revenue generated from a specific can be calculated as follows.

$$REV_h = \alpha \times X_h + \beta \times Y_h - \gamma \times Z_h + \delta \times W_h \tag{1.1}$$

In Fig. 1.3, the expected revenue for a Cloud provider from a specific data center is shown. The data center used is assumed to have a capacity of 1000 VM instances.

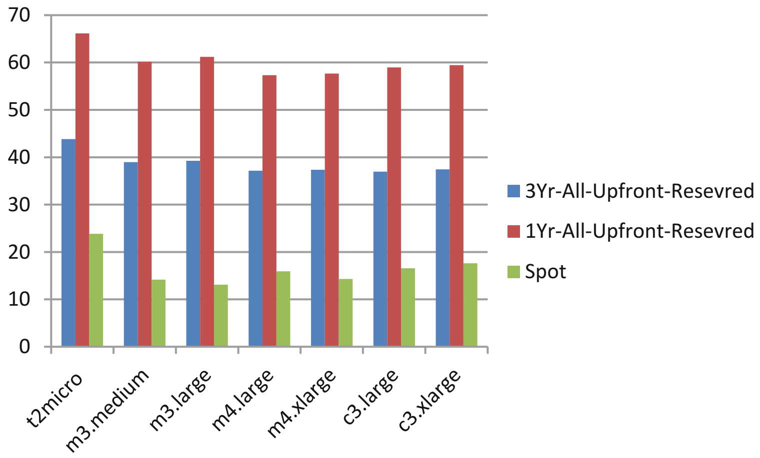


Fig. 1.2 A comparison of Cloud service prices per hour relative to the on-demand pay-as-you-go charging model for a select set of instances offered by Amazon [7]

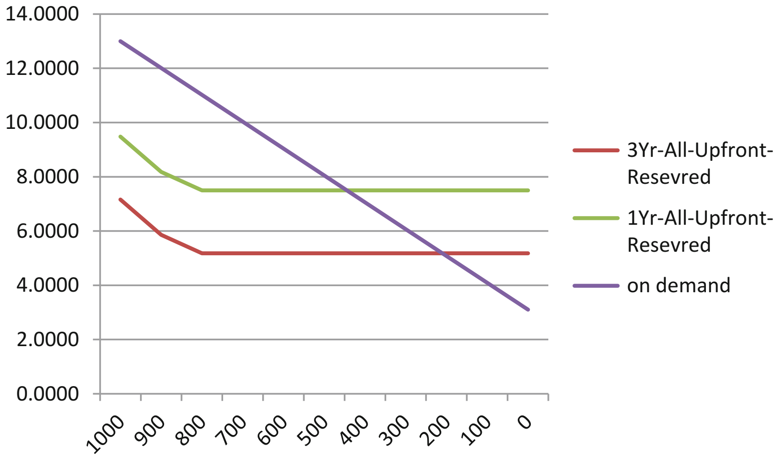


Fig. 1.3 The total revenue achieved by a Cloud of 1000 VM instance capacity under different client demands

The figure represents the expected revenue for a Cloud provider when different types of clients rent instances of the t2.micro type. For each client, revenue is calculated when the demand ranges from 1000 (full Cloud capacity) until 0. In the case of the reserved usage conditions (3 year and 1 year), the original contract is assumed to be 800 instances with any increase rented by the client on an on-demand condition. This figure tells us that with the scattered on-demand pattern which best describes the personal Cloud requests, Cloud providers end up having to auction the remaining capacity as spot instances. This could make it harder for the provider to reach the revenue level achieved when clients use reserved mode. In the figure, the utilization level where the direction/trend changes is (according to this scenario of posted prices) the point where 45% of the data center requests being on-demand. Otherwise, it is evidently better for the provider to reserve the data center resource for the less expensive yet more stable reserved usage condition clients. These are mostly enterprises. This all add extra emphasis on the Cloud providers traffic prediction and resource reservation policies in the case of heavy personal demand. More focus on this area would aid the providers in mitigating the effect of other factors impacting the pricing/commercial side (competitors, client global demand, saturation levels, government-related factors in terms of regulation and taxation).

1.2.2 Energy Efficiency Ramifications

Power consumption in Cloud data centers is a pressing issue for Cloud providers. Power costs represent between 25% and 40% of the operational expenses of a data center [47]. The Natural Resources Defense Council (NDRC) published a data center efficiency assessment in Aug 2014 as an attempt to depict the scale of data centers the world over [36]. The study mentions that US data centers are on track to consume roughly 140 billion kilowatt-hours of electricity annually by 2020, equivalent to the output of 50 large power plants (each with 500 megawatts capacity). If worldwide data centers were a country, they would be the globe's 12th largest consumer of electricity. Another fact here is their assessment of energy efficiency. Their analysis finds that up to 30% of servers are obsolete or not needed and that other machines are grossly underutilized. Persistent issues obscuring efficiency include: peak provisioning, failure to power down unused servers, and competing priorities, keeping costs low and maintaining high levels of security, reliability, and uptime for their clients. These are some of the main factors cited that affect power efficiency and stand in the way of a staggering 40% potential improvement in power consumption. These factors are all largely affected by data center load planning and management. An efficient scheduling energy-aware algorithm would exploit the benefits that come from virtualization technologies, optimal demand-driven provisioning, and efficient load modeling.

In the era of personal Cloud, different energy efficiency algorithms will have added level of complexity to tackle [2, 13, 15]. As discussed in detail in [1], energy efficiency algorithms take one of three routes: server consolidation [10, 24, 42] and switching off of idle server, dynamic voltage and frequency scaling (DVFS) of servers with low loads [33], and algorithms dependent on dynamic prediction [29]. Each of those is affected by the type of load and client distribution change that personal Cloud client request imposes. With consolidation-based solutions, a more heterogeneous set of instances is favorable to counter challenges with network bottlenecks and fault tolerance. DVFS methods could achieve more savings

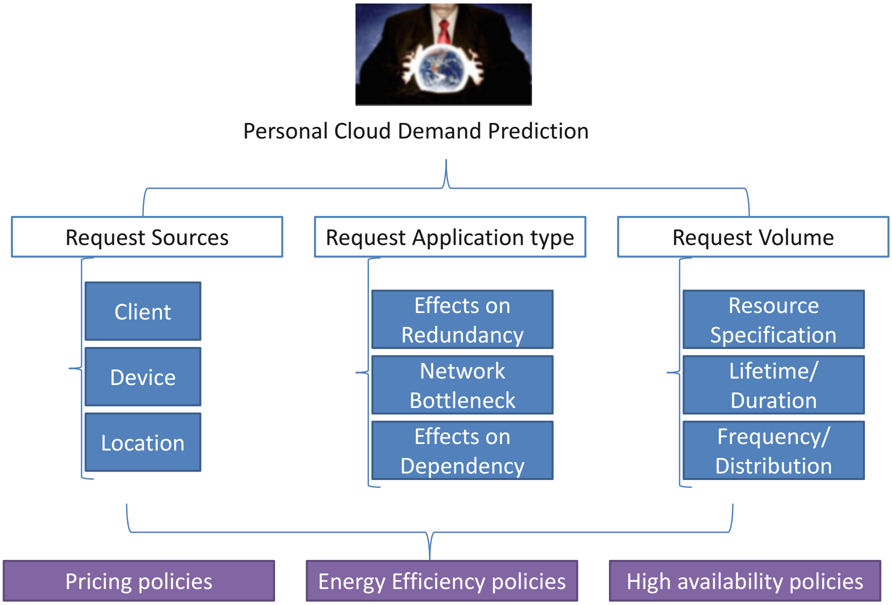


Fig. 1.4 Personal Cloud demand prediction

by exploiting the personal Cloud instances with low activity. Moreover, energy efficiency methods based on future request prediction could benefit from studying personal Cloud load distribution and sizes.

Figure 1.4 summarizes the discussion on personal Cloud client behavior prediction. The upper layer shows the defining elements of the client requests. This includes knowledge about request source in terms of which client, from which device, and from which location. These identifiers would affect the quality of service conditions and can be used to offer customized and personalized service to clients. Next, the application the clients access or use is central to the client behavior modeling. The application type will affect VM instance on redundancy, network bottlenecks and network resource configuration, and on application dependencies. Such factors need to be taken into consideration. The volume of the requests is the third aspect of this triangle. The request volume includes distributions of the resource specification, lifetime or duration of each request, and the probabilistic distribution that better represents arrival rate of requests. Considering the mentioned aspects is critical for Cloud providers when planning the resource allocation policies, energy efficiency policies, high availability policies, and service-pricing policies.

1.3 Request Pattern Prediction

As previously discussed, demand prediction gains elevated criticality in the case of personal Cloud clients. This fact motivated multiple efforts to delve deep into that challenge [3, 5, 12, 28].

In [37], a Cloud workload management framework is introduced which exploits demand forecast, predictive resource allocation, and quality assurance, as well as resource pricing as interdependent components. This framework performs trace analysis of a production video-on-demand (VoD) system. The prediction technique is based on forming a time series to predict video bandwidth demand from online monitoring and determine bandwidth reservations from multiple data centers. An optimization model and study based on game theory is offered to aid in the pricing process.

In [8], the authors propose a Cloud client prediction model evaluated using machine learning techniques including support vector machines (SVM), neural networks (NN), and linear regression (LR). The model covers metrics including response time and throughput. Results and subsequent analysis from the experimentation carried out on Amazon Elastic Compute Cloud (EC2) show that support vector machine provides the best prediction results for random workload traffic pattern.

In [21], a model based on stepwise nonnegative least squares regression is used to predict the demand for a specific Cloud service (viz., HP ePrint service).

A different approach is offered in [14] which is to predict workloads based on identifying similar past occurrences of the current short-term workload history. This process uses a string matching technique. It is based on the assumption that the usage of a Cloud client can sometimes have a repetitive behavior. "Current usage patterns of online services have a probability of having already occurred in the past in a very similar" form[14]. Therefore, it can be used to infer what the future system usage will be for a Cloud client.

Another leading effort can be found in [32]. An investigation of repeatable workload patterns by exploring cross-VM workload correlations coming from application dependency is introduced. A co-clustering is then proposed, based on sample data time series, a co-clustering technique to determine which VMs exhibit high frequency of correlations and the times these are exhibited. This is used in a Hidden Markov Model (HMM) that defines periodical correlations used for workload patterns prediction. Workload behavior prediction will remain an active and ever-growing topic for Cloud providers especially in light of the era of the personal Cloud.

1.4 Internet of Things: The Undiscriminating Swarm of Requests

1.4.1 IoT and Personal Cloud: How Personal Things Are Connected

To show how central Internet of Things implementation tiers are to the personal Cloud, it is enough to look at the Internet Architecture Board (IAB) definition of IoT in RFC 7452.33 "Architectural Considerations in Smart Object Networking": "The term 'Internet of Things' (IoT) denotes a trend where a large number of embedded devices employ communication services offered by the Internet protocols. Many of these devices, often called 'smart objects', are not directly operated by humans, but exist as components in buildings or vehicles, or are spread out in the environment."

These smart objects will represent the elements of the personal Cloud. The personal Cloud of a specific user contains not only the traditional computation/communication devices (laptop, tablet, mobile phone), but it will be mostly represented by other smart objects repressing other devices owned by the client like the ones mentioned in the definition. This would range from fridges alerting clients when grocery items are missing to sensors used to get insights on traffic to nodes used in farms.

To further understand how this works, we look at the communication models envisioned for the IoT and personal Cloud devices.

Devices can communicate in a device-to-device fashion where devices exchange data between each other either through the Internet or other IP networks or through direct connections using Bluetooth or ZigBee protocols, for example. Another type of communication is device to Cloud [4, 9, 20, 31, 39, 50, 51].

"In a device-to-Cloud communication model, the IoT device connects directly to an Internet Cloud service like an application service provider to exchange data and control message traffic"[27]. This happens mostly through Wi-Fi connections. The third model is device-to-gateway model. In the device-to-gateway model, the IoT device connects through an application layer gateway service in order to reach a Cloud service.

1.4.2 Inherent IoT Issues in the Era of Personal Cloud

1.4.2.1 Volume

The sheer numbers of the IoT devices take the probability of risk to new levels. This volume also means that a specific security vulnerability in one device will be magnified as the tendency is to build and use high numbers of identical devices. A poor design for a device or a vulnerability in a communication protocol could have magnified impacts on the Cloud this device is dealing with.

The high volume of small cheap numerous devices means the manufacturers find it harder to break even if they used more sophisticated processors or chips with more memory. This limits the security functionalities/capabilities of a single thing.

1.4.2.2 Interconnectivity

The number of interconnections between things and connections to the Cloud is certainly reaching a new level. IBM expected in [26] that 18.6 billion network connections will exist in 2016. This number along with the things' ability to establish independent and dynamic connections with other devices on their own raises the bar for security tools. It is correct that some of these connections can be neglected or just turned off when necessary, but this is not always the situation. With the increasing dependency, things collecting weather of traffic data cannot be treated the same way we treat household things.

1.4.2.3 The Black Box Syndrome

"Researchers at the French technology institute Eurecom downloaded some 32,000 firmware images from potential IoT device manufacturers and discovered 38 vulnerabilities across 123 products including poor encryption and back doors that could allow unauthorized access"[40]. The cost of manufacturing some of the sensors and other devices constructing the Internet of Things imposes some restrictions in terms of the ability to update or improve these devices. In addition, the implementation visibility of such devices could be close to none."This creates a security vulnerability when a user believes an IoT device is performing certain functions, when in reality it might be performing unwanted functions or collecting more data than the user intends."

1.4.2.4 Privacy Considerations and Legal Issues

The challenge here is the privacy of data collected by IoT devices is susceptible for many reasons. First, the clients agreeing to data being collected is not as straightforward as it was previously. The nature and frequency of data collected or sent by the thing makes it harder for the client to get involved. Second, some IoT devices do not commend an interface to configure privacy settings. Moreover, the data collection and processing can always affect people and devices not involved with the IoT ownership. Examples are people sharing the same geographic location temporarily or sharing the same computing resources. Additional privacy protection measures are required for every carrier of smart devices because as the time goes on there is more of an assumption that each piece of data that can be collected will be collected by some device around you [27, 49].

1.4.3 The Cloud-Friendly Device Market as a Driver of the Personal Cloud

The notebook PC vs. tablet shipment and forecast report in [44] reveals a prime driver behind the personal Cloud. The notebook PC share decreased from 75% vs. 25% in 2011 to 27% vs. 73% for the tablet PC in 2015. Substantial changes are happening. This does not account smartphones which at a fraction of the PC weight are offering very similar video, audio, and processing capabilities to PCs a few years ago. Alcatel-Lucent note the major factors affecting the evolution of the mobile market and the close affiliation with the Cloud [6]. One factor cited by Alcatel is the users' aspiration to operate personal Clouds. This was paired with the increasing demand for Wi-Fi and increasing demand for video conversation in business environments.

1.5 Cloud Brokering Services

Another challenge Cloud providers need to be aware of is the general trend of depending on Cloud service brokers (CSBs) [48]. CSBs serve as the middle layer between Cloud clients and providers. CSBs are divided into two types or flavors: technical CSBs and business ones. Technical CSBs provide an automation, management, or governance fabric underneath which sit one or more Cloud infrastructure products. Business CSBs, on the other hand, are primarily involved with providing a trading interface between Cloud vendors and customers [48].

This does not only apply to enterprises. CSBs can act as Cloud service providers to the public as well. Many brokers are prominent Cloud providers (even in the list of 100 biggest Cloud service providers) including Appirio, BlueWolf, CloudMore, Cloud Nation, Clouditalia, Cloud Sherpas, and Comcast Upware. The roles a CSB can provide to personal Cloud clients in the future can range from intermediary roles like identity or access management capabilities to aggregation of multiple service to insure portability and interoperability in addition to typical service like offering multiple pricing and resource configuration option for personal Cloud clients to choose from.

This is an area that needs attention from Cloud providers. This is evidenced by the overall global Cloud brokerage market growth from $1.57 billion in 2013 to $10.5 billion by 2018 [25, 30, 46]. CSB service evolution will support seamless migration between Clouds (and providers) and will limit data lock-in [11, 19]. The Cloud providers' response to this development will be a deciding factor if this CSB-supported client ability will be a threat or an opportunity.

1.6 Conclusions and Avenues for Future Improvement

The key take away from the chapter is that there is a wide space for improvement for such a new paradigm and for how Cloud providers handle it. With the amount of PC and tablet shipments in the world almost doubling from 2008 to 2015 reaching 545 million devices in 2015, there is definitely a space for growth [38].

The report in [45] gives us an insight into this potential. The report surveys the population's usage/interest in using public services through a Cloud medium. The interest is really high for countries still migrating to the Cloud including India (81% of the sampled population either using government service or interested in using them), UAE (80%), and Brazil (79%). However, these numbers are significantly lower for the UK, the USA, and Germany (59%, 55%, and 53%, respectively). There is some way for the personal Cloud transformation to be complete in the minds of the individual users. Providers need to investigate the areas where performance or service offerings cause that gap in expectation and address these issues. One attempt at figuring this out is in [43]. Within the set of users using the Cloud storage services in the UK in 2014, 55% faced problems while storing or sharing files. These problems included slow speed of access or use, technical server problems, incompatibility and portability problem between devices or file formats, vendor lock-in, privacy challenges with disclosure of data to third party, or unauthorized use of information by service provider. The last one is an issue for the public with doubts over agreements and government usage/ownership of data very much in the public's mind. Guarantees for data access and movement within the Cloud are expected by the client. This is really consistent with the results from [41] which alluded that ease of accessing files from several devices or locations was a main reason for using the Cloud. "Among the reasons for using Cloud services, three in five users of Internet storage space identified the possibility to use files from several devices or locations and to share files with other persons easily (59%). Furthermore, over half of users gave protection against data loss as a reason for using the Cloud (55%). Users also valued having a larger memory space (44%) and being able to access large libraries of music, films, or TV programs (22%)."

Concerns about security and privacy remain a main barrier to the use of Cloud services.

For the cases where people were aware of Cloud services but chose not to use them, "44% cited security or privacy concerns as a reason for not making use of such services, 28% mentioned concerns about the reliability of service providers and 22% cited a lack of skills" [41]. These insights draw a path for Cloud providers in order to increase the personal Cloud penetration and achieve higher client satisfaction.

References

1. M. Abu Sharkh, A. Shami, An Evergreen Cloud: Optimizing Energy Efficiency in Green Cloud Computing Environments Based on Virtual Machine State Prediction, Submitted to Elsevier computer networks journal, Mar 2016.
2. M. Abu Sharkh, A. Shami, An evergreen cloud: optimizing energy efficiency in heterogeneous cloud computing architectures. Veh. Commun. **9**, 199–210 (2017)
3. M. Abu Sharkh, M. Jammal, A. Ouda, A. Shami, Resource allocation in a network-based cloud computing environment: design challenges. Commun. Mag. IEEE **51**(11), 46–52 (2013)
4. M. Abu Sharkh, A. Kanso, A. Shami, P. Öhlén, Building a cloud on Earth: a study of cloud computing data center simulators. Comput. Netw. J. **108**, 78–96 (2015)
5. M. Abu Sharkh, A. Shami, A. Ouda, Optimal and suboptimal resource allocation techniques in Cloud computing data centers. J. Cloud Comput. Springer **6**(1), 6 (2017) [Online]. Available: https://link.springer.com/article/10.1186/s13677-017-0075-2
6. Alcatel Lucent, Mobile conversations in the era of the personal Cloud, Application note [Online]. Available: http://www.tmcnet.com/tmc/whitepapers/documents/whitepapers/2013/8057-alcatel-lucent-mobile-conversations-the-era-the-personal.pdf
7. Amazon, Amazon EC2 Pricing page [Online]. Available: https://aws.amazon.com/ec2/spot/pricing/
8. A.Bankole Cloud client prediction models for cloud resource provisioning in a multitier web application environment, An M.Sc. thesis Department of Systems and Computer Engineering, Carleton University, Ottawa, Apr 2013
9. E. Baccarelli, P.G.V. Naranjo, M. Scarpiniti, M. Shojafar, J.H. Abawajy, Fog of everything: energy-efficient networked computing architectures research challenges and a case study. Access IEEE **5**, 9882–9910 (2017). ISSN:2169-3536
10. A. Beloglazov, R. Buyya, Energy efficient resource management in virtualized Cloud data centers, in *2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*
11. L.F. Bittencourt, M. Moraes, I. Petri, O.F. Rana, Towards virtual machine migration in fog computing, in *2015 10th International Conference on P2P Parallel Grid Cloud and Internet Computing (3PGCIC)*, 2015, pp. 1–8
12. F. Bonomi, R. Milito, P. Natarajan, J. Zhu, Fog computing: a platform for Internet of things and analytics, in *Big Data and Internet of Things: A Roadmap for Smart Environments* (Springer, Cham, 2014), pp. 169–186
13. D. Bruneo, A. Lhoas, F. Longo, A. Puliafito, Modeling and evaluation of energy policies in green clouds. IEEE Trans. Parallel Distrib. Syst. **26**(11), 3052–3065 (2015)
14. E. Caron, F. Desprez, A. Muresan, Forecasting for grid and cloud computing on-demand resources based on pattern matching, in *2nd IEEE International Conference on Cloud Computing Technology and Science*, Nov 2010, pp 456–463
15. B.G. Chun et al., An energy case for hybrid datacenters. ACM SIGOPS Oper. Sys. Rev. **44**(1), 76–80 (2010)
16. Cisco, Cisco global cloud index: forecast and methodology, 2014–2019 white paper. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-Cloud-index-gci/Cloud_Index_White_Paper.html
17. L. Colombus, IDC Predicts SaaS Enterprise Applications Will Be A $50.8B Market By 2018 [Online]. Available: https://www.forbes.com/sites/louiscolumbus/2014/12/20/idc-predicts-saas-enterprise-applications-will-be-a-50-8b-market-by-2018/#2657e46222a8. Accessed Feb 2018
18. L. Columbus, Roundup of small and medium business cloud computing forecasts and market estimates. [Online] Available: https://www.forbes.com/sites/louiscolumbus/2017/04/29/roundup-of-cloud-computing-forecasts-2017/#1dfc95c731e8. Last accessed 27 Aug 2018

19. B. Darrow, is-live-migration-coming-to-amazon-web-services-smart-money-says-yes/. Available: https://gigaom.com/2014/10/12/is-live-migration-coming-to-amazon-web-services-smart-money-says-yes/
20. R. Deng, R. Lu, C. Lai, T. H. Luan, H. Liang, Optimal workload allocation in fog-Cloud computing toward balanced delay and power consumption. Internet Things J. IEEE **3**, 1171–1181 (2016). ISSN:2327-4662
21. V. Garg, L. Cherkasova, S. Packirisami, J. Rolia, Workload analysis and demand prediction for the HP ePrint service in *2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, May 2013, pp. 1054–1059
22. Gartner, Inc, Gartner says the personal Cloud will replace the personal computer as the center of users' digital lives by 2014 [Online]. Available: http://www.gartner.com/newsroom/id/1947315
23. Gartner, Inc, Gartner says 8.4 bILLION cONNECTED tHINGS wILL bE IN use in 2017, Up 31 pERCENT from 2016, Gartner Press Release [Online]. available: https://www.gartner.com/newsroom/id/3598917
24. H. Goudarzi, M. Pedram, Energy-efficient virtual machine replication and placement in a Cloud computing system, in *Proceedings of IEEE Cloud*, Jun 2012, pp. 750–757
25. N. Henderson, 2015 Talkin' Cloud 100: Top Cloud Services Providers [Online]. Available: http://talkinCloud.com/tc100
26. IBM's The big data and analytics hub, The 4 Vs of the Big Data, IBM, [Online]. Available: http://www.ibmbigdatahub.com/infographic/four-vs.-big-data
27. Internet Society Organization, The Internet of things, an overview, (Oct 2010) [Online]. Available: https://www.internetsociety.org/sites/default/files/ISOC-IoT-Overview-20151014_0.pdf
28. M. Jammal, A. Kanso, A. Shami, High availability-aware optimization digest for applications deployment in cloud. in *Proceedings of ICC*, 2015, in press. Available: http://vixra.org/pdf/1410.0193v1.pdf
29. G. Jung, M.A. Hiltunen, K.R. Joshi, R.D. Schlichting, C. Pu, Mistral: dynamically managing power, performance, and adaptation cost in cloud infrastructures, in *Proceedings of IEEE 30th International Conference on Distributed Computing Systems (ICDCS)*, 2010
30. B. Kepes, New stats from the state Of Cloud report [Online]. Available: http://www.forbes.com/sites/benkepes/2015/03/04/new-stats-from-the-state-of-Cloud-report/#23bbd67326f9
31. D. Kelaidonis, A. Rouskas, V. Stavroulaki, P. Demestichas, P. Vlacheas, A federated edge Cloud-IoT architecture, in *2016 European Conference on Networks and Communications (EuCNC)*, 2016, pp. 230–234
32. A. Khan, X. Yan, S. Tao, N. Anerousis, Workload characterization and prediction in the cloud: a multiple time series approach, in *2012 IEEE Network Operations and Management Symposium (NOMS)*, Apr 2012, pp. 1287–1294.
33. D. Kliazovich, P. Bouvry, S.U. Khan, GreenCloud: a packet-level simulator of energy-aware Cloud computing data centers. J. Supercomput. Spec. Issue Green Netw. **62**(3), 1263–1283 (2012)
34. A. Kohgadai, 12 must know statistics on Cloud usage in the enterprise [Online]. Available: https://www.skyhighnetworks.com/Cloud-security-blog/12-must-know-statistics-on-Cloud-usage-in-the-enterprise/
35. P. Lamson, A look at U.S. & European Cloud adoption trends & forming lasting partnerships [Online]. Available: https://gtdc.org/evs/Carbonite_USvsEU.pdf. Accessed Oct 2015
36. National Resources Defense Council, America's data centers are wasting huge amounts of energy [Online]. Available: http://www.nrdc.org/energy/files/data-center-efficiency-assessment-IB.pdf. Accessed Oct 2015
37. D. Niu, Demand forecast, resource allocation and pricing for multimedia delivery from the cloud, Ph.D. thesis, Department of Electrical and Computer Engineering, University of Toronto, Canada, May 2013

38. NPD DisplaySearch, Quarterly mobile PC shipment and forecast report (Jan 2013), [Online]. Available: http://www.tmcnet.com/tmc/whitepapers/documents/whitepapers/2013/8057-alcatel-lucent-mobile-conversations-the-era-the-personal.pdf
39. F. Al-Turjman, Fog-based caching in software-defined information-centric networks. Comput. Electr. Eng. J. **69**(1), 54–67 (2018). Elsevier
40. R. Samani, 3 key security challenges for the Internet of things [Online]. Available: http://www.securingtomorrow.com/blog/knowledge/3-key-security-challenges-internet-things/
41. H. Seybert, P. Reinecke, Internet and Cloud services – statistics on the use by individuals: Half of Europeans used the Internet on the go and a fifth saved files on Internet storage space in 2014 [Online]. Available: http://ec.europa.eu/eurostat/statistics-explained/index.php/Internet_and_Cloud_services_-_statistics_on_the_use_by_individuals
42. S. Srikantaiah, A. Kansal, F. Zhao, Energy aware consolidation for Cloud computing, in *Proceedings of the 2008 Conference on Power Aware Computing and Systems*, pp. 10–16
43. Statista Team, Problems experienced when saving or sharing files on the internet in Great Britain in 2014 [Online]. Available: http://www.statista.com/statistics/321310/problems-with-online-file-saving-or-sharing-great-britain-uk/
44. Statista Team, PC, tablet and Internet/Cloud-computing systems unit shipments worldwide from 2002 to 2015 (in million units) [Online]. Available: http://www.statista.com/statistics/282549/worldwide-pc-and-tablet-shipments/
45. Statista Team, Would you personally like to use Cloud computing in the future when interacting with government departments offering public services? [Online]. Available: http://www.statista.com/statistics/384604/Cloud-computing-public-service-usage/
46. Talking Cloud team, Cloud Services Brokerage Company List and FAQ [Online]. Available: http://talkinCloud.com/Cloud-services-broker/Cloud-services-brokerage-company-list-and-faq
47. C. Tulkoff, Using immersion cooling technology for enhanced efficiency & reliability for data center servers [Online]. Available: http://www.slideshare.net/CherylTulkoff/df-r-webinar-immersion-cooling-technology. Accessed Oct 2015
48. K. Weins, Cloud Computing Trends: 2015 State of the Cloud Survey [Online]. Available: http://www.rightscale.com/blog/Cloud-industry-insights/Cloud-computing-trends-2015-state-Cloud-survey. Accessed Oct 2015
49. Wind River, Security in the Internet of things lessons from the past for the connected future [Online]. Available: http://www.windriver.com/whitepapers/security-in-the-internet-of-things/wr_security-in-the-internet-of-things.pdf
50. W. Zhang, Z. Zhang, H. Chao, Cooperative fog computing for dealing with big data in the internet of vehicles: architecture and hierarchical resource management. Commun. Mag. IEEE **55**, 60–67 (2017). ISSN:0163-6804
51. J. Zhu, D.S. Chan, M.S. Prabhu, P. Natarajan, H. Hu, F. Bonomi, Improving web sites performance using edge servers in fog computing architecture, in *2013 IEEE Seventh International Symposium on Service-Oriented System Engineering*, Redwood City, 2013, pp. 320–323. https://doi.org/10.1109/SOSE.2013.73

Chapter 2
Optimization in Edge Computing and Small-Cell Networks

Jitender Grover and Ram Murthy Garimella

2.1 Introduction

2.1.1 Edge Computing

Distributed computing paradigms like Cluster Computing and Grid Computing were innovated to meet the computing need arising in various applications. The culmination of such efforts is a Cloud Computing paradigm enabled by data centers, and distributed across the world and the Internet. It is commercially very successful.

Presently, IoT devices are connected to the Cloud through some (wired/wireless) communication medium. IoT-Cloud infrastructure is shown in Fig. 2.1. Sensed data reaches to the Cloud via core Internet for processing. This leads to high communication cost (in terms of time delay), which can be catastrophic for many hard real-time delay-sensitive applications such as health monitoring, automated vehicles, manufacturing systems/assembly lines, video surveillance, etc. (Fig. 2.2).

It was realized that in the context of technologies like IoT and CPS, the delay involved in processing data using Cloud resources will be high. It is necessary to perform computing, communication/networking, and control (actuation) locally, leading to the idea of Edge Computing [1]. Variations of Edge Computing led to paradigms such as Fog/Mist/Dew computing as shown in Fig. 2.3. It can be considered as a hierarchical Cloud where Dew has least latency and least processing power and Cloud has the highest latency and highest processing power.

J. Grover (✉) · R. M. Garimella
Signal Processing and Communication Research Center (SPCRC), International Institute of Information Technology, Hyderabad, Telangana, India
e-mail: jitender.grover@research.iiit.ac.in; rammurthy@iiit.ac.in

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_2

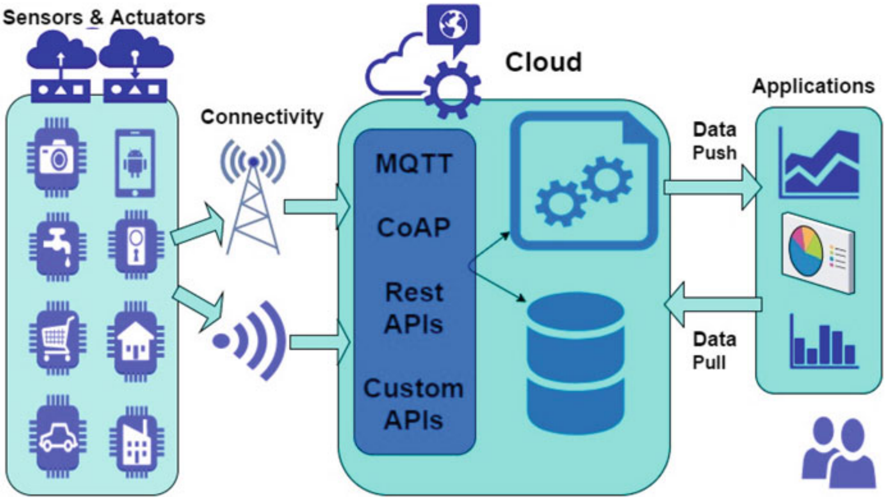


Fig. 2.1 IoT-Cloud infrastructure

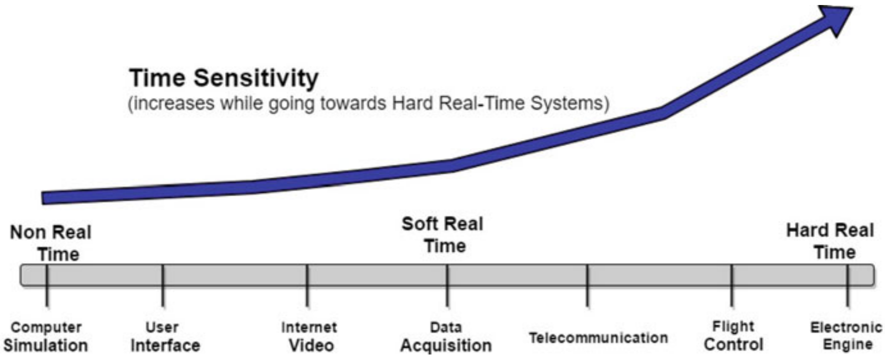


Fig. 2.2 Time sensitivity in real-time vs non-real-time systems

Edge Computing potentially addresses issues like latency concerns, limited processing/storage capabilities of devices/things, battery life, network bandwidth constraints, security and privacy concerns, etc.

Edge computing [2] paradigm enables allocation of computing resources to the tasks generated by IoT applications. Some tasks are highly delay constrained. Also, the cost of Edge Computing infrastructure needs to be minimized [3]. Since the connectivity structure of Edge Computing infrastructure is under the control of the user, the grid-based architecture provides a good approximation. Thus, we are naturally led to an interesting joint optimization problem with cost, delay, etc. as the objective functions.

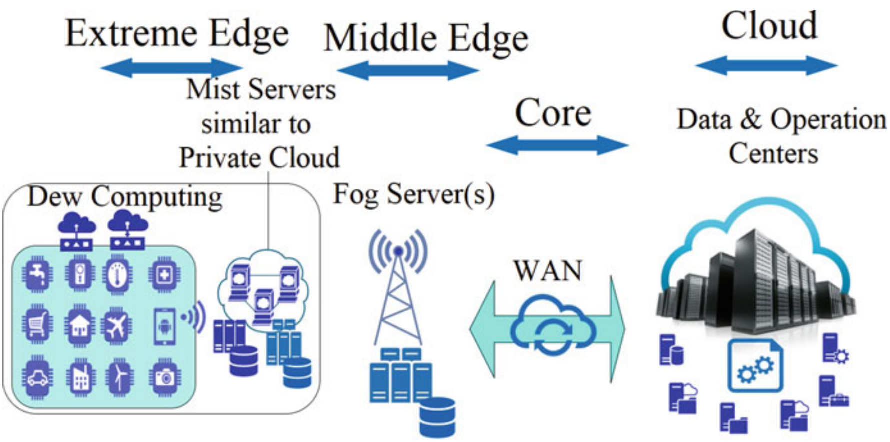


Fig. 2.3 Cloud hierarchy

2.1.2 Edge Communication

It is most likely that the edge wireless network is based on micro/pico/femto cell infrastructure (Fig. 2.4). A multi-tier communication network usually overlays multiple tiers of cells and potentially shares a common spectrum. A macro-cell is usually a traditional cell tower which covers 15–30 kms, but small cells cover a smaller area with respect to macro-cell. A micro-cell covers 1–2 kms. They can be used temporarily during the large events or in heavy rush areas for additional coverage. Pico-cell can manage up to 100 users and covers 250 m. They are mostly installed indoor and used to improve the coverage in an office or shopping area. Femto-cells are often self-installed and can manage a few users only. Small cells increase frequency reuse multiple times and enable cellular networks to manage high number of IoT devices on the edge. They can be installed at homes, offices, etc. to provide seamless connectivity to all the available IoT devices.

In most interesting Edge Computing applications (such as in IoT), the packet traffic on the network has predictable patterns. Hence dynamic channel allocation based on historical traffic data ensures optimal utilization of available channels.

It is natural to employ cognitive radio technology to increase the spectrum utilization (as the demand for channels is sporadic by devices). To facilitate cognition, a wideband spectrum sensing needs to be done in a time-optimal manner. Thus, we are led to a design of optimal cellular, cognitive radio network (femto/pico cell).

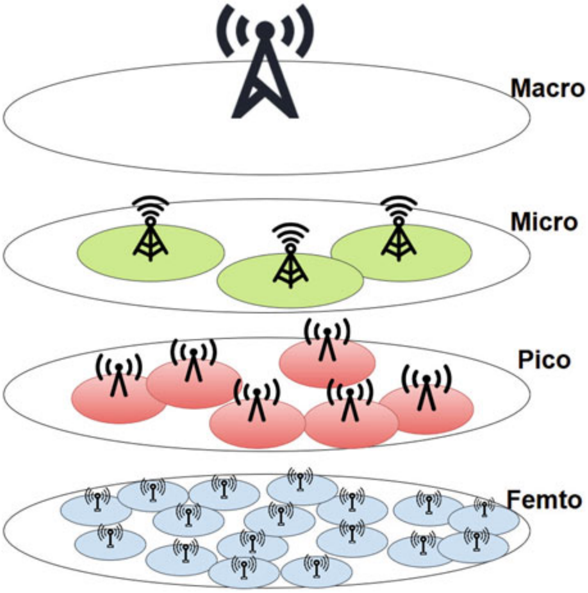


Fig. 2.4 Multi-tier communication networks

2.1.3 Edge Application

From the above discussion, the need for Edge Computing and Edge Communication is clearly evident. We now focus our attention on one application as an example which highlights finer requirements of the Edge Computing infrastructure.

Hospital is one of the applications which requires Edge Computing infrastructure to deal with delay-sensitive patient monitoring system. It should be kept in mind that critically ill patients effectively require real-time processing of the diagnostic data and real-time intervention by the doctor. Some regular patients generate diagnostic data that leads to predictable demands of edge resources (computation/communication). But some type of data generated by patients in unpredictable. Hence, the Edge Computing paradigm leads to flexibility in provisioning edge resources based on demand.

In most hospitals, various type of data (e.g., diagnostic data) with varied processing time constraints are generated. For instance, the data has features of Big Data. On the other hand, storage of medical records of past and present patients needs to be done. The processing of such medical records to mine interesting patterns could be done using Cloud Computing resources. Thus, Cloud Computing and Edge Computing complement each other. Effectively real-time/non-real-time processing of patient data effectively aids medical diagnosis.

The Edge Computing paradigm effectively has the following characteristics needed by various applications.

- Upgradability of computing/communication, control resources on demand.
- QoS provisioning for various applications.
- Security/privacy of data processed, etc.

It is crystal clear that there is a need for Edge Computing paradigms (such as Fog, Mist, Dew computing). The resources at the edge are mainly computational resources (e.g., multi-core processor-based systems), communication/networking resources, and control/actuation resources. These resources must be shared in an optimal manner to meet the QoS constraints of applications.

In this chapter, we formulate and solve interesting optimization problems arising in the design of Edge Computing and Edge Communication network suitable for many applications. The motivation behind Edge Computing Optimization problem is to minimize the cost of edge servers' deployment, taking location, type, and cost of servers into consideration. In practical scenario, if an industry tries to provide Edge Computing solution and wants to decrease the deployment cost, the proposed optimization solution can be used. Motivation behind small-cell optimization problem is to allocate channels dynamically to meet the future demands of IoT devices. The use of micro-cells, pico-cells, and even femto-cells is very much possible in the near future to provide seamless connectivity to IoT devices. Optimization of multilevel dynamic channel allocation (DCA) can help ISPs to cover more number of devices efficiently and provide better performance to the end users.

2.2 Related Work

This section briefly describes the efforts made by the researchers to explain the need of Edge Computing. Further, the optimization strategies proposed for the edge servers' deployment would be discussed in this section. Amin et al. [4] discussed the participatory Edge Computing for the local community services. They used and proved that using the local server is better if we have a complex core network available. They successfully implemented and provided third-party applications on local servers for the community near to them. Kim et al. [5] worked with IoT devices to offload their computation on the Cloud. Based on the results, authors found it better to offload and compute on the Cloud rather than doing computation on the IoT device itself. Samie et al. [6] described a way of distributed computation offloading to many remote machines for distributed QoS for IoTs. It suggests to decompose the problem into small processes among many resources. Dynamic programming is used for resource allocation, and the whole problem is solved using ILP. Mao et al. [7] went one step further and offloaded the computation data to the edge server which is nearer as compared to the Cloud.

By many examples it has been proved by the researchers that even though the machines on the edge have lesser configuration, still the delay is higher while getting the computation results from the Cloud. It happens due to the communication cost. In [8], the author proposed the optimizing solution by formulating it as an integer nonlinear program for offloading and resource allocation in Mobile Edge Computing. The author found that due to the hardness of the problem, optimally solving it for the last scale network is impractical. To solve it efficiently the problem is divided into resource allocation and task offloading. Resource allocation problem is further divided into two subproblems which are uplink power allocation and computing resource allocation. Convex and quasi-convex optimization techniques are used to solve these problems.

Now, this section briefly describes a few cases where researchers tried to optimize the edge server deployment strategy. Qiang Fan et al. [9] gave a strategy CAPABLE to optimize the cost of server deployment and end-to-end delay between client devices and resources. The simulation results show that the scheme can trade off between deployment cost and delay. Qin et al. [10] have proposed a software-defined approach to manage heterogeneous IoT and sensor devices. This is done via providing the best matching resource for different classes of IoT devices. Authors have utilized a reflective middleware with a layered IoT SDN controller for managing various IoT applications. It is the extension of multi-network information architecture (MINA). Farah Slim et al. [11] proposes a mechanism related to the multidimensional Cloud. It gives an analytical model for blocking analysis. It also tried to find out the best strategy for distributed edge placement. Further, a strategy is devised for resource allocation and capacity planning on the edge network. So authors tried to optimize the limited resources available at the edge network.

For achieving energy efficiency after taking feedback latency into account, mobile units demand for a common/different Cloud server for processing [12]. The proposed algorithm gives a good performance for IoT with multi-small-cell Edge Computing and MIMO. The problem is formulated as the minimization problem but doesn't use learning methods to optimize the device-edge combinations. The proposed framework also works for radio access point implemented in distributed and parallel manner with limited signaling to the Cloud. It is compared with disjoint optimization algorithm and shows better results. In IoT-Edge computing scenario, it is very much necessary to learn the best combinations over the period of time.

This section covers some of the related proposed solutions, but it is very clear that very few researchers are focusing on methods for optimizing the allocation of edge computation and communication resource to the IoT devices for their applications. Further in the chapter, we have shown how Integer Linear Programming can be used for optimal allocation of edge resources.

2.3 Edge Computing Architecture: Integer Linear Programming Formulation

Consider devices/things which are locally distributed in space with limited local computation/memory resources (e.g., nodes in the wireless sensor network). They require high-end Edge Computing platforms. Also, it is necessary to be able to schedule the computing tasks onto a relatively high-end Edge Computing server. The goal is to minimize the cost as well as the delay in processing the tasks required by local devices/things. We now provide the detailed modeling assumptions.

- There are two types $\{TypeA, TypeB\}$ of Edge Computing devices with costs $\{C_A, C_B\}$ (model can easily be generalized to the finite number of types of high-end computing devices).
- The devices/things are distributed in a rectangular grid (model can be generalized to an arbitrary graph connectivity of devices)

The objective is to minimize the total cost of high-end computing platforms (required to process the "delay sensitive" tasks submitted by the devices/things) while at the same time ensuring that the things distributed on the rectangular grid are able to schedule their tasks onto a certain minimum number of high-end Edge Computing platforms. We now formulate the optimization problem as a {0, 1} (binary/integer) linear programming problem.

Let the variable associated with placement of high-end Edge Computing platforms on the rectangular grid be $\{a_{ij}, b_{ij}$ where $0 \le i \le M, 0 \le j \le N\}$, i.e.:

$$\begin{aligned} a_{ij} &= 1;\text{ if Type A computer is placed at (i, j)} \\ &\qquad\qquad \text{location on the rectangular grid} \\ &= 0;\text{ otherwise} \end{aligned} \tag{2.1}$$

$$\begin{aligned} b_{ij} &= 1;\text{ if Type B computer is placed at (i, j)} \\ &\qquad\qquad \text{location on rectangular grid} \\ &= 0;\text{ otherwise} \end{aligned} \tag{2.2}$$

Let the grid points on the rectangular grid (where devices/things and edge computers are located), i.e., (M)(N) points, be divided into mutually disjoint sets $\{D_\gamma : 1 \le \gamma \le L\}$, i.e., those sets constitute a set partition of all points on the rectangular grid. The constraints are that points in each set D_γ (for all j) are served by at least "S" high-end Edge Computing platforms. Thus the optimization problem has a linear objective function

$$\sum_{i=1}^{M}\sum_{j=1}^{N} C_A a_{ij} + \sum_{i=1}^{M}\sum_{j=1}^{N} C_B b_{ij} \tag{2.3}$$

subject to the linear constraints

$$\sum_{(i,j)\in D_K} a_{ij} + \sum_{(i,j)\in D_K} b_{ij} \ge S,\ \ for\ 1 \le K \le L \tag{2.4}$$

This integer ({0, 1}) linear programming problem is solved using well-known techniques. Efficient algorithms exist for solving the problem.

An alternate formulation requires that every set D_δ (in the set partition of (M)(N) grid points) is covered by at least S_1, type A computers and by at least S_2, type B computers. Thus, the constraints in the above linear programming problem (Eq. 2.4) get modified in the following manner:

$$\left.\begin{array}{l}\sum_{(i,j)\in D_K} a_{ij} \geq S_1;\\ \sum_{(i,j)\in D_K} b_{ij} \geq S_2;\end{array}\right\} \text{for } 1 \leq K \leq L \tag{2.5}$$

We now discuss the high-end Edge Computing allocation scheme for different IoT applications.

2.4 Optimized Resource Allocation at Edge

After the deployment of the edge servers, next issue that is required to be addressed is allocation of different types of edge servers to different types of applications according to their requirements. There can be many types of tasks taken care of by the edge servers like image classification, audio/video classification, sensed data fusion, data analytics, mathematical calculations for output generation, etc. These tasks can be related to IoT applications such as healthcare, environment monitoring, urban security, smart water/electricity management system, wearable, automated transportation, smart surveillance, etc.

Now considering that there are two or more types of edge servers (high and low cost) available on the edge, allocation of appropriate server to different applications is necessary to complete them in the best possible ways. This can be done by allocating cost A servers to one type of tasks and cost B servers to other type of tasks and so on (depending upon the type of servers available). The division of tasks can be majorly dependent upon the following:

- strict delay requirements of the task
- reliability requirement
- frequency of requests generated by an application and some other negligible factors.

For example, image classification, smart surveillance, automated transportation system, etc. related tasks should be assigned to an edge server with more powerful resources, near to the source to reduce delay, with less probability to failure, and able to manage frequent continuous requests. On the other hand, tasks like mathematical calculations, function-based outputs, etc. can be given to the low-cost edge servers. In the first case, delay more than threshold or failure in response can be catastrophic. Based on this, optimization function can be updated with more constraints.

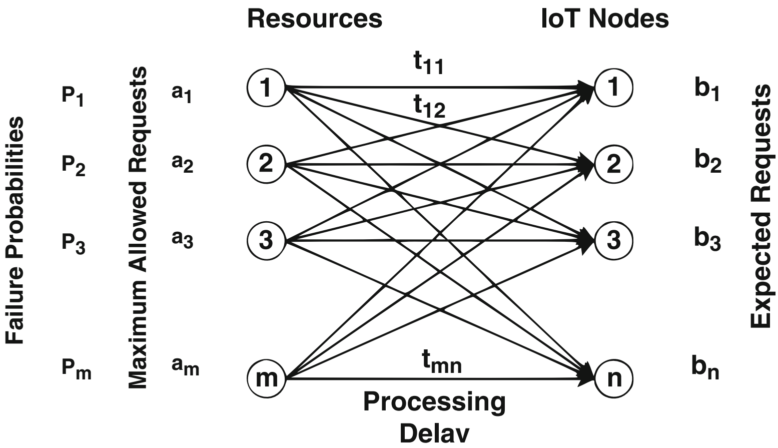


Fig. 2.5 Resource allocation problem

As shown in Fig. 2.5:

- The maximum capacity of m edge servers is represented by the vector $[a_1, a_2, \ldots, a_m]$
- Probability of Failure of these systems is shown by the vector $[p_1, p_2, \ldots, p_m]$
- The expected number of requests from n IoT nodes is displayed as vector $[b_1, b_2, \ldots, b_n]$
- A matrix $[t_{ij}]$ $\forall\ i = 1, 2, \ldots, m$ & $j = 1, 2, \ldots, n$ is used to map IoT nodes to edge servers as resources

Now, a dual-objective resource allocation optimization problem has been formed as per Fig. 2.5. Considering cost-effective deployment of edge servers and allocation of servers to the IoT devices in an appropriate way, Eq. (2.3) is not good enough. If x_{ij} be the number of allocated requests to a service proving node i from an IoT node j, multi-objective optimization problem is formulated as the following:

$$\begin{aligned} &\text{minimize} \sum_{j=1}^{n} t_{ij} x_{ij}\ \forall\ 1 \le i \le m \\ &\text{and} \\ &\text{minimize}\ \boldsymbol{\pi}_{i=1}^{m} p_i^{\sum_{j=1}^{n} x_{ij}} \end{aligned} \tag{2.6}$$

subject to constraints

$$\sum_{j=1}^{n} x_{ij} \le a_i\ \forall\ i = 1, 2, \ldots, m \tag{2.7}$$

$$\sum_{i=1}^{m} x_{ij} = b_i \ \forall \ j = 1, 2, \ldots, n \tag{2.8}$$

$$x_{ij} \geq 0 \ \forall \ i = 1, 2, \ldots, m \& j = 1, 2, \ldots, n \tag{2.9}$$

Equation (2.6) is the updated objective function for resource allocation. This multi-objective function is trying to minimize the cost of deployment, total time taken for the processing of all the available requests, and the failure probability associated with all the edge servers. Equation (2.7) keeps a check on the number of requests to be sent to a server that should not exceed its capacity. Equation (2.8) divides the load from an IoT node to edge servers. Equation (2.9) puts a lower bound on the allocations. The following condition must be satisfied to achieve feasible solution:

$$\sum_{i=1}^{m} a_i >= \sum_{j=1}^{n} b_j \tag{2.10}$$

In case of not meeting the condition given in Eq. (2.10), more edge servers are required to be added to satisfy the demand. The edge server can transfer the requests to other neighbor edge servers as well in that situation. In worst scenario, if no such servers are available, then the request can be transferred to Cloud as well.

2.5 Optimal Edge Computing Communication Network: Integer Linear Programming

It is expected that small cells (pico/femto), enabling efficient frequency reuse, will be the important innovation in the deployment of the 5G cellular network. We consider the case where the wireless communication network providing connectivity between devices/things at the edge is infrastructure based.

The goal is to make efficient utilization of channels available in pico/femto cell. Thus, we are naturally led to dynamic channel allocation based on historical traffic data (at the edge) in adjacent small cells. Further, we propose dynamic spectrum access based on time-optimal spectrum sensing (using CR approach). We formulate these approaches as optimization problems. We reason below that these are two related optimization problems. In [13], the time-optimal spectrum sensing problem was formulated and solved. We utilize the solution for efficient spectrum sensing in small cells. We now formulate and solve the dynamic channel allocation problem in adjacent small cells and propose a solution.

Let there be M adjacent cells. Also, let the historical traffic (on some time unit example hour) in the spectrum bands in those cells be $\{n_1, n_2, n_3, \ldots, n_M\}$. We normalize $n_i' s$ to arrive at the probability mass function, i.e.:

$$p_i = \frac{n_i}{\sum_{j=1}^{M} n_j}; \text{ for } 1 \le i \le M \tag{2.11}$$

IDEA: In those bands where "p_i" is small, allocate small number of channels, i.e., $n_i's$. On the contrary if p_i is allocated higher number of channels in such a way that

$$\sum_{i=1}^{M} n_i.p_i = E[Z] \tag{2.12}$$

subject to the constraint that

$$\sum_{i=1}^{M} n_i = L \tag{2.13}$$

i.e., average number of channels allocated to a small cell are maximized. It should be noted that "Z" is the random variable associated with number of channels allocated per cell.

We reason below that, if there are no constraints imposed on $n_i's$, the problem becomes trivial.

Problem formulation: Maximize $E[Z]$

Subject to the constraint that

$$\sum_{i=1}^{M} n_i = L \tag{2.14}$$

where L is the total number of available channels.

For instance, we can order the probabilities in the order

$$q_1 \le q_2 \le \ldots \le q_M \tag{2.15}$$

- Set $n_M = L$ and $n_i = 0$ for $i \ne M$. Thus, L is the maximum attainable value for $E[Z]$. Thus, we are naturally led to imposing reasonable constraints on $n_i's$ motivated by practical considerations.
- $n_i's$ are in A.P., i.e., $n_1, n_2 = n_1 + d, n_3 = n_1 + 2d, \ldots, n_M = n_1 + (M-1)d$

From the point of view of allocation of channels, these constraints are very reasonable and implementable.

Now, we solve the above precise optimization problem. We consider the case where the number of channels allocated per cell are in Arithmetic Progression. Thus, the constraint leads to

$$n_1 + (n_1 + d) + \ldots + (n_1 + (M-1)d) = L \tag{2.16}$$

Since $\{n_1, d\}$ are integers, we are led to the linear Diophantine equation

$$2Mn_1 + d(M)(M-1) = 2L \tag{2.17}$$

Thus, we would like to utilize solutions of the above Eq. (2.17) and solve the stochastic optimization problem of maximizing $E[Z]$.

We first sort the probabilities, that is, $p_i^{'}s$, in increasing order, resulting in relabeled probabilities $q_i^{'}s$, that is:

$$q_1 \leq q_2 \leq \ldots \leq q_M \tag{2.18}$$

Thus, maximization of $E[Z]$ requires that

$$\widetilde{n}_1 \leq \widetilde{n}_2 \leq \ldots \leq \widetilde{n}_M \tag{2.19}$$

It can be readily seen that

$$E[Z] = \widetilde{n}_1 + (\delta)d; \text{ where } \delta = \sum_{j=1}^{M}(j-1)q_j \tag{2.20}$$

Also, computing variance of Z, we have that

$$Var[Z] = (\alpha - \delta^2)d^2; \text{ where } \alpha = \sum_{j=1}^{M}(j-1)^2 q_j \tag{2.21}$$

The following theorem provides unique solution to the above optimization problem.

Theorem *Assume that* $\{(a_1, d_1), \ldots, (a_l, d_l), \ldots, (a_k, d_k)\}$ *are the set of solutions of linear Diophantine Eq. (2.17) among the infinitely many solutions, which are positive real integers. If* $a_1 < \ldots < a_l < \ldots < a_k$*, then* $d_1 > \ldots > d_l > \ldots > d_k$*. In such a case,* (a_1, d_1) *is the best solution, which maximizes the expected value* $E(Y)$*.*

Proof To prove that (a_1, d_1) maximizes $E(Y)$, we need to prove the below expression

$$a_l + \delta d_l \leq a_1 + \delta d_1 \qquad \text{for } l = 2, 3, \ldots, K \tag{2.22}$$

$$\text{i.e. } a_l + \delta d_l \leq a_1 + \delta d_1 \qquad \text{for } l = 2, 3, \ldots, K \tag{2.23}$$

$$\text{i.e. } \frac{a_l - a_1}{d_1 - d_l} \leq \delta \qquad \text{for } l = 2, 3, \ldots, K \tag{2.24}$$

Since, (a_1, d_1) and (a_l, d_l), both satisfy the linear Diophantine equation (2.17). Therefore, we can write

$$M \times a_1 + \frac{(M-1)M}{2} \times d_1 = L \tag{2.25}$$

$$M \times a_1 = L - \frac{(M-1)M}{2} \times d_1 \tag{2.26}$$

$$M \times a_l + \frac{(M-1)M}{2} \times d_l = L \tag{2.27}$$

$$M \times a_l = L - \frac{(M-1)M}{2} \times d_l \tag{2.28}$$

Multiplying Eq. (2.24) by M on both sides, we get

$$M \times \frac{a_l - a_1}{d_1 - d_l} \leq M \times \delta \tag{2.29}$$

$$\frac{M \times a_l - M \times a_1}{(d_1 - d_l)} \leq M \times \delta \tag{2.30}$$

$$\frac{[L - \frac{(M-1)M}{2} \times d_l] - [L - \frac{(M-1)M}{2} \times d_1]}{(d_1 - d_l)} \leq M \times \delta \tag{2.31}$$

$$\frac{\frac{(M-1)M}{2} \times (d_1 - d_l)}{(d_1 - d_l)} \leq M \times \delta \tag{2.32}$$

$$\frac{(M-1)}{2} \leq \delta \tag{2.33}$$

Substituting the value of δ, we get

$$\frac{(M-1)}{2} \leq \sum_{j=1}^{M} (j-1) p_j \tag{2.34}$$

Therefore, it is enough to prove the above expression in order to prove (a_1, d_1) is the best solution which maximizes $E(Y)$.

By the proof of contradiction, it can be proved that the equality in Eq. (2.34) holds only for uniform distribution, i.e., $p_1 = p_2 = \ldots = p_M = \frac{1}{M}$. Therefore,

$$\sum_{j=1}^{M}(j-1)p_j = \frac{1}{M}\sum_{j=1}^{M}(j-1) = \frac{(M-1)}{2} \tag{2.35}$$

For any other probability distribution, we get

$$\frac{(M-1)}{2} < \sum_{j=1}^{M}(j-1)p_j \tag{2.36}$$

Proof To prove the minimum value of $\sum_{j=1}^{M}(j-1)p_j$ occurs only for uniform distribution, i.e., $p_1 = p_2 = \ldots = p_M = \frac{1}{M}$, we use the proof of contradiction. Suppose that the probability masses $p_1 = q - \epsilon$, $p_2 = p_3 = \ldots = p_{M-1} = q = \frac{1}{M}$ and $p_M = q + \epsilon$ gives the minimum value of $\sum_{j=1}^{M}(j-1)p_j$, where $\epsilon > 0$ is a small positive real number. Using the given probabilities, we can calculate value of the expression $\sum_{j=1}^{M}(j-1)p_j$ as below:

$$\begin{aligned}\sum_{j=1}^{M}(j-1)p_j = 0 \times (q-\epsilon) + \sum_{j=1}^{M}(j-1) \times q \\ + (M-1) \times (q+\epsilon)\end{aligned} \tag{2.37}$$

$$\begin{aligned}\sum_{j=1}^{M}(j-1)p_j &= \sum_{j=1}^{M}(j-1) \times q + (M-1) \times (q+\epsilon) \\ &= \sum_{j=1}^{M}(j-1) \times q + (M-1) \times \epsilon \\ &= \frac{1}{M} \times \sum_{j=1}^{M}(j-1) + (M-1) \times \epsilon \\ &= \frac{1}{M} \times \frac{(M-1)M}{2} + (M-1) \times \epsilon \\ &= \frac{(M-1)}{2} + (M-1) \times \epsilon\end{aligned} \tag{2.38}$$

Since, $\epsilon > 0$, it contradicts our supposition. Therefore, we can conclude that the value of $\sum_{j=1}^{M}(j-1)p_j > \frac{(M-1)}{2}$ for any probability distribution and minimum value occurs for uniform distribution.

2.6 Conclusion

In this book chapter, we have solved the cost-effective edge server placement problem optimally, considering the location of the edge servers are grid points. This is solved using two types of edge servers that can easily be generalized to any number of servers' type. Efficient allocation of resources is also taken into consideration, and this issue is also solved using another optimization function. Along with this, we also solved the issue of optimal spectrum sensing arising in cellular cognitive radio networks. This book chapter also solves the issue of optimal dynamic channel allocation problem arising in small cells based on historical traffic data. In future, cost effective edge placement optimization problem can be solved based on a graph where the location of the edge servers can be arbitrary.

References

1. M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, The case for VM-based cloudlets in mobile computing. IEEE Pervasive Comput. **8**(4), 14–23 (2009)
2. M. Satyanarayanan, The emergence of edge computing. Computer **50**(1), 30–39 (2017)
3. S. Sardellitti, G. Scutari, S. Barbarossa, Joint optimization of radio and computational resources for multicell mobile-edge computing. IEEE Trans. Signal Inf. Process. over Netw. **1**(2), 89–103 (2015)
4. A.M. Khan, F. Freitag, On participatory service provision at the network edge with community home gateways. Proc. Comput. Sci. **109**, 311–318 (2017)
5. S. Kim, Nested game-based computation offloading scheme for mobile cloud iot systems. EURASIP J. Wirel. Commun. Netw. **2015**(1), 229 (2015)
6. F. Samie, V. Tsoutsouras, S. Xydis, L. Bauer, D. Soudris, J. Henkel. Distributed QoS management for internet of things under resource constraints, in *2016 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ ISSS)* (IEEE, 2016), pp. 1–10
7. Y. Mao, J. Zhang, K.B. Letaief, Dynamic computation offloading for mobile-edge computing with energy harvesting devices. IEEE J. Sel. Areas Commun. **34**(12), 3590–3605 (2016)
8. T.X. Tran, & D. Pompili, Joint task offloading and resource allocation for multi-server mobile-edge computing networks. arXiv preprint arXiv:1705.00704 (2017)
9. Q. Fan, N. Ansari. Cost aware cloudlet placement for big data processing at the edge, in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6
10. Z. Qin, G. Denker, C. Giannelli, P. Bellavista, N. Venkatasubramanian. A software defined networking architecture for the internet-of-things, in *Network Operations and Management Symposium (NOMS), 2014 IEEE* (IEEE, 2014), pp. 1–9
11. F. Slim, F. Guillemin, Y. Hadjadj-Aoul, On virtual network functions' placement in future distributed edge cloud, in *2017 IEEE 6th International Conference on Cloud Networking (CloudNet)*, Sept 2017, pp. 1–4
12. S. Sardellitti, G. Scutari, S. Barbarossa, Joint optimization of radio and computational resources for multicell mobile-edge computing. IEEE Trans. Signal Inf. Process. Over Netw. **1**(2), 89–103 (2015)
13. G.R. Murthy, R.P. Singh, S. Abhijeet, S. Chandhary, Time optimal spectrum sensing. arXiv preprint arXiv:1606.02849 (2016)

Chapter 3
A Comprehensive Survey on Architecture for Big Data Processing in Mobile Edge Computing Environments

Maninder Jeet Kaur

3.1 Introduction

Mobile networks have evolved significantly over the past few decades from 2G to 3G and further from 3G to 4G LTE. This growing mobile traffic is driven by consumer smart phones, streaming video messaging, and P2P applications. In addition, this growth is expected to increase more as businesses extend their processes to smart mobile devices after the evolution of mobile cloud computing and services such as m-learning, m-gaming, m-healthcare, m-commerce, m-governance, etc., are directly accessible from the mobile devices. This also implies that large amount of data is involved and high network load and growing demand of network bandwidth as data have to be transmitted and received to and from the mobile devices and cloud data centers [1]. With existing centralized cloud computing architecture, the applications are facing harsh challenges. For example, mobile devices connected to the distant centralized cloud servers try to obtain high computing applications which enforce additional load both on the radio access networks and on the backhaul networks and, thus, introducing high latency [2]. As estimated by Cisco, the number of devices connected to the IoT will be around 50 billion by 2020, and at the same time, mobile data traffic is predicted to continue doubling each year [3, 4]. However, the load if this large volume of data which is generated by end users and processed at cloud data centers leads to certain challenges like stringent latency, capacity constraints, resource-constrained devices, uninterrupted services with intermittent connectivity, and enhanced security [5].

There are two emerging paradigms which have been proposed to help overcome these challenges: (i) cloud-based radio access network (CRAN) that aims at the

M. J. Kaur (✉)
Department of Engineering, Amity University Dubai, Dubai International Academic City, Dubai, UAE
e-mail: mkaur@amityuniversity.ae

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_3

centralization of base station (BS) functions via virtualization and (ii) mobile edge computing (MEC) which proposes to empower the network edge [6]. CRAN brings cloud computing technologies into mobile networks by centralizing baseband processing units (BBU) of radio access network to ensure highly efficient network operation and flexible service delivery when handling mobile internet traffic surging [7]. CRAN uses centralized BBU to do baseband processing, while MEC handles distributed task offloading by shifting computation capacity from a public cloud to an edge cloud, which significantly reduces latency. In MEC, network edge devices, such as base stations (BSs), access points, and routers, are endowed with, albeit limited, computing and storage capabilities to serve users' requests as a substitute of clouds, while significantly reducing the transmission latency as they are placed in the proximity of end users.

3.2 Mobile Edge Computing: Overview and Framework

The first real-world MEC platform was introduced by Nokia Networks in 2013. In that, the computing platform called radio applications cloud servers (RACS) – was fully integrated with the flexi multiradio base station [19]. Later the term, mobile edge computing was standardized by European Telecommunications Standards Institute (ETSI) and Industry Specifications Group (ISG). It is acknowledged by European 5G PPP (5G Infrastructure Public Private Partnership) as a prime emerging technology for 5G networks [8].

The ETSI Mobile Edge Computing Industry Specification Group (MEC ISG) was formed in December 2014. Its aim was "to create a standardized, open environment which will allow the efficient and seamless integration of applications from vendors, service providers, and third-parties across multi-vendor mobile edge computing platforms" [9]. The ISG recently received a 2-year extension – beginning in March 2017. The ISG modified its terms of reference to include the investigation of the use of edge computing beyond a "mobile-only" access environment such as areas with Wi-Fi and fixed access technologies. Therefore, ISG will be changing its name from "mobile edge computing" to "multiaccess edge computing" to reflect the change in terms of reference, while still maintaining the widely recognized "MEC" acronym [10]. 3GPP has also recently published their work plan toward 5G and more specifically toward meeting IMT-2020 requirements, which includes study items and development of technical reports on scenarios that drive 5G architecture and interfaces. Major standardization bodies, ITU (International Telecommunication Union), and European Telecommunications Standards Institute (ETSI) have initiated activities relating to 5G [11–14] with commercial deployments expected in 2020.

MEC enables cloud computing capabilities at the edge of the cellular network. It also enables the applications and related processing tasks like content caching and media processing to run on the edge of the network, i.e., more closer to the mobile subscribers. This not only improves the user experience by providing high availability of the content with reduced latency, but also puts control of the service delivery in the hands of mobile network operator (MNO). This further allows the MNO to manage user experience and also helps to rapidly deploy new applications and services to the subscribers [15]. MEC can be implemented within the existing 3GPP mobile network infrastructure. Therefore, it is expected to be integral part of 5G specifications.

The cloud service platforms will be providing the computation and other services to the mobile communication networks with the help of huge alternative architecture centralized data centers which are the backbone of the network. A signal or data packet usually spends RTT (round trip time) of several milliseconds to travel between the mobile device and the cloud in order to be processed in the data center which might be several kilometers away. To reduce the RTT to the range of less than 1 ms, the distance between the mobile and cloud can be no greater than 100 km. But the closest places or “hops” to a user are the base stations at the edge of the network which can be used to outsource the cloud closer to the user providing low latency. Therefore, in response to the above requirement, the mobile operators are working on mobile edge computing (MEC) in which the computing, storage, and networking resources are integrated with the base stations. Hence, all compute intensive and latency sensitive applications like augmented reality and image processing can be hosted at the edge of the network. This provides the end user with swift and powerful computing, energy efficiency, storage capacity, mobility, location, and context awareness support [16, 17]. A decentralized architecture with compute and storage capability at the edge of the network provides an intermediate processing stage to reduce the amount of data shipped back to the cloud by executing algorithms for applications such as face recognition, load building, and landmark labeling for augmented reality (AR) – a cognitive assistance or even cloud-sourced video analytics. The communications service providers (CSPs) will need to respond fast to capacity increases while maintaining latency.

According to the white paper published by ETSI, mobile edge computing can be characterized by [18]:

- On Premises: MEC platforms can run isolated from the rest of the network, while they have access to local resources. This is very important for machine-to-machine scenarios. The MEC property of segregation from other networks also makes it less vulnerable.
- Proximity: Being deployed at the nearest location, mobile edge computing has an advantage to analyze and materialize big data. It is also beneficial for compute-hungry devices such as segmented reality, video analytics, etc.
- Lower latency: Mobile edge computing services are deployed at the nearest location to user devices, isolating network data movement from the core network. Hence, user experience is accounted high quality with ultra-low latency and high bandwidth.
- Location awareness: Edge-distributed devices utilize low-level signaling for information sharing. MEC receives information from edge devices within the local access network to discover the location of services.

- Network context information: Applications providing network information and services of real-time network data can benefit businesses and events by implementing MEC in their business model. On the basis of RAN real-time information, these applications can estimate the congestion of the radio cell and network bandwidth. This will help them in future to make smart decisions for better delivery of services to customers.

In this chapter, an extensive survey of the mobile edge computing in the field of big data processing is presented along with architecture, framework, applications, and challenges. Finally, deep learning approach has been discussed in the context of MEC applications. Deep learning has drawn a lot of academic and industrial interest and has been applied to various classification tasks, natural language processing, object detection, etc.

3.2.1 Framework

A general view of basic cellular network is shown in Fig. 3.1. The core network is wire connected (e.g., IP/ethernet) with radio access network (RAN). RAN connects the base station with backhaul network through the interface that supports high data transfer rate [19]. Wireless user equipment connects through RAN to the mobile operator network.

RAN facilitates the connection between mobile phones and the mobile core network. It also covers the wide geographical area divided into several cells, and each cell is integrated with its base station. Base stations are further connected to the base station controller (BSC) via microwave or WLL. It is also known as radio network controller (RNC) which is responsible to control the base station node and

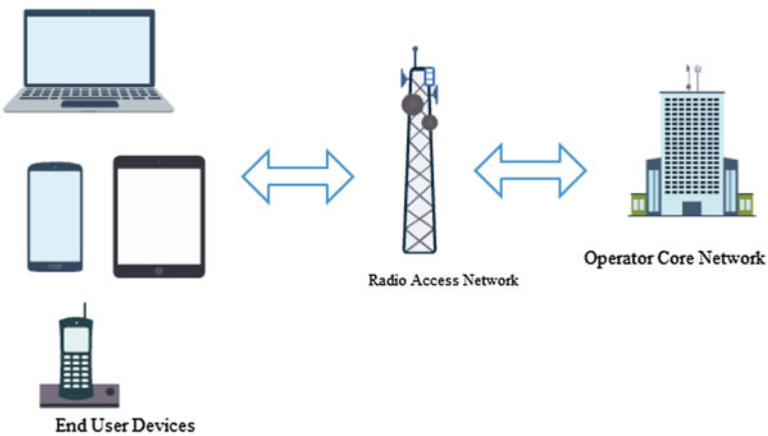


Fig. 3.1 Basic cellular architecture

carry out the mobility management functions. In the framework of 4G long-term evolution and future 5G mobile network architecture, different solutions have been proposed to overcome the challenges which involve capacity and latency of current radio access network (RAN). With an accustomed support of mobile internet and digital radio signaling capability using time division multiple access (TDMA), LTE 4G got an edge over the other wireless telecommunication technologies providing the best QoE.

A centralized RAN system, also known as cloud-RAN (C-RAN), concentrates different processing resources together to form a pool in a central data center. This not only reduces deployment costs but also leverages low latency connections between different RAN processing units enabling a series of enhanced capabilities. Using C-RAN, the number of cell sites will be reduced, and the user will be offered better QoE while maintaining similar coverage and reducing operating expenses. Figure 3.2 shows the general architecture of mobile edge computing. The MEC layer resides between the mobile devices and the cloud. It mostly compiles with cloud computing to support and enhance the performance of the mobile subscribers [20–23]. As shown, the mobile edge servers are computing equipment installed at or near base stations. Unlike centralized cloud servers or peer-to-peer mobile devices, MEC is managed locally by the network operator. The generic computing resources within the mobile edge servers are virtualized and are exposed via application program interfaces (APIs), so that they are accessible by both user and operator applications. Also, these provide local virtual machines (VMs) to serve the computation needs of mobile devices often with much lower latency than remote cloud servers. They also serve some other functions such as content caching, traffic monitoring, local information aggregation, and user location services.

A simplified framework architecture for MEC is shown in Fig. 3.3. It depicts the high-level functional entities that are further grouped in the system level, host level,

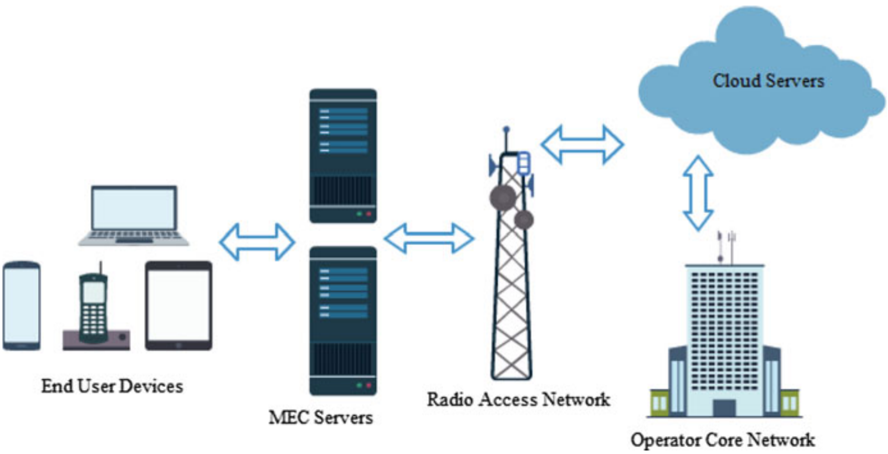


Fig. 3.2 Mobile edge computing (MEC) architecture

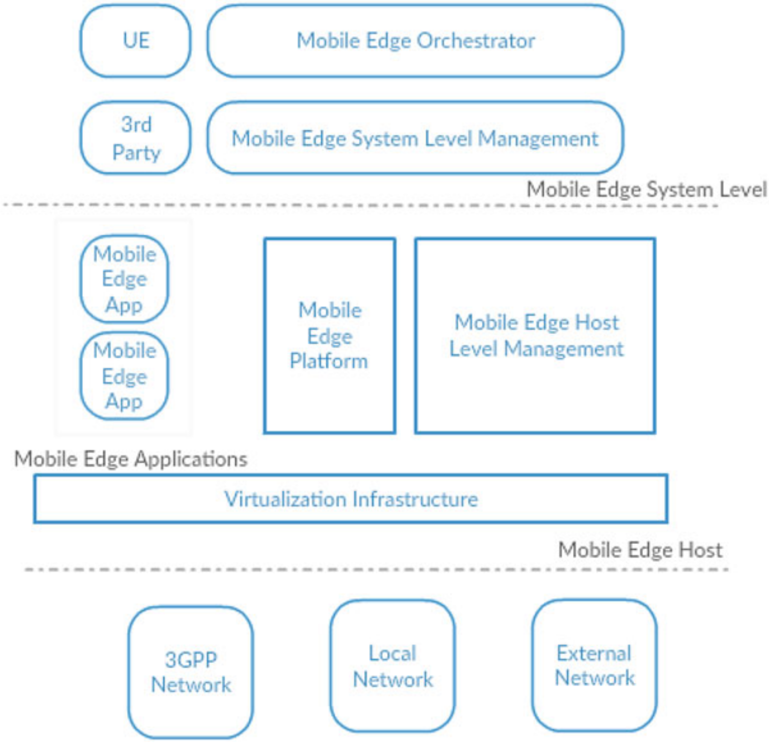


Fig. 3.3 Mobile edge computing (MEC) framework

and network level. At the mobile edge system level, the *mobile edge orchestrator* serves the central role of coordinating among the UEs, the mobile edge hosts, and the network operator. It records accounting and topological information about the deployed mobile edge hosts, available resources, and available mobile edge services. Also, it interfaces with the *virtualization infrastructure* and maintains authentication and validation of application packages. It maintains information on the whole mobile edge system about all the deployed hosts, the services and resources available in each host, the applications that are instantiated, and the topology of the network. It manages the applications by checking the integrity and authenticity of the application, validating the policies for the applications, and maintaining a catalog of the applications that are available.

At the mobile edge host level, *mobile edge applications* run VMs supported by *virtualization infrastructure* within the mobile edge host. They provide services like computational job execution and UE location information. The virtualization infrastructure includes a data plane that executes the forwarding rules received by the mobile edge platform and routes the traffic among the applications, services, and the networks. It also manages the virtualized resources for the mobile edge applications, which consists of allocating and releasing virtualized computation, storage, and network resources. It provides support for fault and performance monitoring by collecting and reporting information on virtualized resources and providing information further to server- and system-level management.

The *mobile edge platform* hosts mobile edge services. It interacts with mobile edge applications, so that they can advertise, discover, offer, and consume mobile edge services. The *mobile edge platform manager* provides element management functions to the mobile edge platform and administers application essentials such as life cycle, service requirements, operational rules, domain name system (DNS) configuration, and security. In addition, mobile edge host can provide persistent storage and time of the day information for the applications. The mobile edge platform receives the traffic forwarding rules from the mobile edge platform manager, mobile edge applications, and mobile edge services and based on those as well as on policies it provides the instructions to the forwarding plane. Mobile edge platform can communicate with other mobile edge platforms over reference point, which is intended for control plane procedures. Using this interface, platforms may be grouped together and can form a communications grid.

3.3 Role of MEC Toward Big Data: Applications and Challenges

IoT devices are expected to generate a significant amount of data as their use becomes ubiquitous. IoT cloud communication models and big data generated by devices result in increased latency and incremental data transfer cost [24–26]. Using cloud computing with the IoT will raise concerns about privacy, and also the most of the IoT end nodes are power constrained. While design considerations are intended to improve the performance, the processing and memory capabilities will be constrained. Offloading some computing tasks to the network edge could be the solution – called mobile edge computing. MEC can address these issues by serving as an aggregation point. MEC consists of geodistributed servers or virtual servers with built-in IT services. These can be implemented locally at mobile user premises. MEC may utilize cellular network elements like base station, Wi-Fi access point, LTE base station, or a 3G/LTE cell aggregation site which can be indoor or outdoor. Therefore, it works on both downstream data on behalf of cloud services and upstream data on behalf of IoT services.

3.3.1 Applications

MEC offers high bandwidth and low latency mobile access to both information and computation resources, because of its proximity to the mobile devices. Hence, reducing the resource demands on the mobile backhaul. The mobile edge host is a privileged location to run applications with sensitive real-time requirement, given its proximity to where relevant data traffic is generated and delivered. By 2018, 40% of the IoT-created data will be stored, processed, analyzed, and acted upon close to or at the edge of the network [27].

- Surveillance and Safety – In applications that require intensive computational load like video surveillance and pattern recognition, MEC with big data processing plays an important role. Real-time automated surveillance requires low communication delays which will be fulfilled by mobile edge host's computing resources. A local network of IoT devices, e.g., smart phones, etc., are connected to the broadband mobile network through local IoT gateway. The video streams are conveyed to the mobile edge host, where the application for video surveillance is running. The real-time processing is also carried out, so that if there is an error in the recorded video, the application can send a trigger to the central office located in the internet. This configuration prevents to deploy costly dedicated equipment and connectivity infrastructure, by simply leveraging the radio access network and mobile edge system from the mobile network operator [28, 29].
- Offloading – With the motivation to increase limited computing, storage, bandwidth, or battery capacities of mobile devices by referring to external, rich resource systems, MEC serves as additional support. The compute intensive tasks are offloaded either because they cannot be executed in-time by the devices due to limited hardware capabilities or to reduce the power consumption of the mobile devices. In case, no MEC server is available, mobile devices can degrade gracefully to a more distant MEC server, internet cloud servers, or their own hardware resources [28, 29]. Examples of such jobs include high-speed browser, 3D rendering, video analysis, sensor data processing, and language translation. With MEC as an option to cheaply offload heavy computation to nearby mobile edge hosts for accelerated processing, the mobile device's computational capability an energy consumption will no longer be the bottleneck in delivering rich applications. This is particularly appealing in the IoT environment, where the mobile devices are likely to consist mostly of small sensors and other equipment with minuscule processors and limited energy supply, even though offloading introduces cost due to higher system complexity
- Augmented reality (AR), gaming – These applications all require low latency. The heavy computation requirements of jobs such as physical simulation and artificial intelligence might overwhelm the limited capability of the devices. Offloading these jobs to a MEC server provides an appropriate balance between computation power and proximity.
- Healthcare – There have been exponential growth in development and deployment of human computer interaction devices recently, such as smart phone, smart watch, etc. To add more reliability to the existing system of cloud-based healthcare, MEC helps health advisors to assist their patients, independent of their geographical location. It enables smartphones to collect patient physiological information like pulse rate, temperature, etc., and send it to the cloud server for storage, data sync, and sharing. Health advisers will have access to the

cloud server, and they can immediately diagnose the patients and can assist them accordingly [30–34].

- Moving IoT devices and connected vehicles – The motivation behind the IoT use cases is being connected to the network and also mobility with connectivity, for example, drones, vehicles, etc. The connection scenario can be vehicle to vehicle, vehicle to access point, or access point to access point. MEC offers local processing and low latency to such situations and is expected to be important to 5G services. In fact, mobility and session management is usually performed by the core network functions, for example, serving gateway (SGW) and serving GPRS support node (SGSN) in the current 3G/4G networks. Mobile edge applications for IoT would not require complex traffic filtering and manipulation in order to receive the desired traffic and the mobile edge platform can leverage on legacy network entities to carry out consistent control plane operations, e.g., for mobility as well as other crucial aspects associated to gating, QoS enforcement, charging, etc. Therefore, MEC enables scalable, reliable, and distributed environments that are synced with the local sensors [35].

3.3.2 Challenges

In this section, we will discuss the challenges for deployment and implementation of real-time data processing at the edge of the network [36, 37]. Current cloud computing frameworks like Amazon Web Service, Microsoft Azure, and Google App Engine can support data-intensive applications but data processing [38–40].

- Threats related to infrastructure – The concept of the edge network computing is part of the "last mile network" which means that it uses many technologies to build a network. While same is advantageous, and at the same time prone to several types of attacks like DoS (Denial of Service) attacks and jamming. These two types of attacks can easily consume bandwidth, frequency band, and computing resources at the edge. Also, there could be rogue users that can intrude the system and perform some malicious activities. For example, a device can be reconfigured and set to send fake information or any incorrect data. Hence, any compromised device connected in a cluster environment can change and control services in that cluster.
- Virtualization – There are various challenges related to virtualization technology – which is based to share the resources in the mobile edge computing environment. Therefore, if one resource is compromised, it can affect the whole virtualization infrastructure. One of the common security concerns is privacy leakage. The APIs implemented to deliver the information should be protected against any kind of malicious activities. The network functions virtualization (NFV) is a core enabling technology employed by the mobile edge hosts to create network elements like routers, packet gateways, and internet protocol (IP) multimedia subsystems, using generic hardware. By separating software from

hardware, it allows dynamic provisioning of services and flexible deployment of network functions. However, the performance of current NFV implementations often falls well below that of dedicated hardware network equipment. This can be managed by small cells, where the base stations are located, and their associated mobile edge hosts would be required to have flexible installation [41, 42].

- Mobility management – Mobility is one of the main aspects of mobile edge computing. Usually when the handoff occurs from one base station to the other, the other active applications running on the mobile hosts and servers should remain intact. Also, the virtual machines (VMs) created on that host must migrate to the new host efficiently. This whole process involves the communication as well as the computation handoff which is quite an expensive and complicated affair itself [43]. Therefore, it will be necessary to accurately model the handoff procedure by facilitating high density of base stations in 5G and also with a trade-off between the frequency and the data rate. Therefore, the main challenge to manage the mobility efficiently is to balance these trade-offs in a complicated scenario.
- Discovering edge nodes and computing on edge nodes – MEC involves the mechanisms to discover the appropriate nodes on the network that can act as edge. Due to the large volume of devices and data involved, the process would need to be rapid enough so as to allow seamless integration of the resources. Hence, it is a challenge to deal with the edge nodes and the edge hosts reliably and proactively, and hence, more efficient methods would be required to make it practical. After the discovery of the edge nodes, another challenge is to incorporate high power computing resources that can handle analytical computing tasks. Also, these resources would be required to be portable across different environments.

3.4 Deep Learning for Reliable and Intelligent Mobile Edge Computing for Big Data Processing Applications: State of the Art

Big data processing systems in mobile environments require ultra-low latency and reliable data analytics solutions that can combine in real-time heterogeneous mix of data stemming from the IoT mobile devices. Such data analytics capabilities cannot be provided by cloud-centric data processing techniques. With higher capability of smart mobile devices, a distributed mobile edge computing architecture can leverage deep learning techniques for *reliable mobile sensing*, *automate the detection process*, and also to *detect the security threats*. In this section, we will survey the state-of-the-art research efforts on the same issue in the context of MEC.

Deep learning is a promising approach for extracting accurate information from IoT devices deployed in complex environments. Deep learning is more suitable for mobile edge computing due to its multilayer structure. Deep neural networks provide highly accurate classifiers which consist of different layers that follow each other. Each layer is a simple function of its previous layer, representing a more sophisticated concept than the previous layer. Therefore, the initial layer is the raw input data, and the final layer gives the result of the inference.

Machine learning methods process the data streams by performing supervised, unsupervised, semisupervised, and deep learning models [44]. The data mining and machine learning methods are useful for early knowledge discovery from big data streams. Therefore, these methods are useful for real-time big data analytics where multiple learning models at different levels of big data systems filter the data streams and uncover the knowledge patterns in parallel [45–47]. Deep neural network (DNN) is the technology behind voice-based personal assistance [48], self-driving cars [49], automatic image processing [50], etc., and finance analytics has quickly adopted machine learning to harness large volume of data in areas of fraud detection, risk management, and compliance. Various smart services are being pushed to edge devices. For example, Intel's Movidius Neural Compute Stick [51] is a tiny deep learning device that one can use to accelerate AI programming and DNN inference application deployment at the edge. Basically, DNNs are the types of artificial neural networks (ANN) that work as function approximators used to analyze the correlation between relevant features and the output of the processed data. Therefore, DNNs are ANN with multiple hidden layers [52].

3.4.1 Deep Learning for Mobile Sensing

Deep neural networks (DNNs) can model high-level abstractions in data due to feature hierarchy that allows each hidden layer to train on a distinct feature set based on the previous hidden layer's output. Hence, DNNs are able to recognize and model complex features from large datasets like video, radar readings from various IoT mobile devices. Various machine learning algorithms can be employed for edge analytics in mobile environment like recurrent neural networks (RNNs), convolutional neural networks (CNNs), restricted Boltzmann machines (RBMs), generative adversarial networks (GANs), and long short-term memory (LSTM) [53–56]. In this particular case of mobile sensing, recurrent neural networks (RNNs) are more suitable for variable length sequential data like the one produced by mobile sensors [57, 58]. Some other deep learning and the related issues are depicted in Table 3.1.

3.4.2 Deep Learning for Automatic Detection Process

In this section, we investigate the overall data processing workflow in collection and analysis of broadcast messages from the base stations in the mobile environment. These broadcast messages are continuously being transmitted from the base stations

Table 3.1 Deep learning methods, description, and issues for mobile sensing

S.no.	Method	Description and issues
1	Residual nets [59]	These introduce shortcut connections into CNNs, which greatly reduces the difficulty of training super deep models. However, since residual nets mainly focus on visual inputs, they lose the capability to model temporal relationships, which are great importance in time-series sensor inputs.
2	LRCN [60]	These apply CNNs to extract features for each video frame and combine video frame sequences with LSTM [61], which exploits spatial-temporal relationships between video inputs. However, it does not consider modeling multimodal inputs. This capability is important to mobile sensing and computing tasks.
3	DBM [62]	Multimodal DBMs merge multimodal inputs such as images and text, with deep Boltzmann machines (DBMs). However, the work does not model temporal relationships and does not apply tailored structures such as CNNs to effectively and efficiently exploit local interactions within input data.
4	RBM [63] and MultiRBM [64]	These use deep Boltzmann machines and multimodal DBMs to improve the performance of heterogeneous human activity recognition.
5	IDNet [65]	It applies CNNs to the biometric gait analysis task.
6	DeepX [66] and RedEye [67]	These reduce the energy consumption of deep neural networks, based on software and hardware, respectively.

LRCN long-term recurrent convolutional networks, *LSTM* long short-term memory, *DBM* deep Boltzmann machines, *RBM* restricted Boltzmann machines, *RNN* recurrent neural networks, *CNN* convolutional neural networks

and dictate how mobile phones interact with them. These messages also include information on how the base station communicates with other base stations and handles cellular traffic in the surrounding RF environment. The broadcast messages are decoded, and the resultant textual data are processed and used as input to the machine learning algorithms [68]. LTE technology, which is based on UMTS, is optimized for all-IP traffic and has several advantages including high network throughput and low latency.

There are many existing methods to detect the base station routers (BSRs) that use global positioning system (GPS) for location information. These methods are time lapse based or configuration based. However, in case of mobile edge computing, flexible and more efficient solutions are required, which can automatically detect the presence of noncommercial BSRs without the prior need of location information, extensive time-lapse surveys or without scanning the databases of known commercial carriers so as to reduce the comeback time. Hence, there is a need for automated data mining techniques to understand the various patterns in big data generated by the mobiles and base stations. Most of the state-of-the-art methods focus on unsupervised clustering which is the most popular data mining method used for finding hidden patterns in data. Table 3.2 gives a brief on some of machine learning methods.

Table 3.2 Various machine learning techniques for automatic detection

S.No.	Method	Description
1	SOM [69–72]	Self-organizing map (SOM) is a type of artificial neural network that is trained using machine learning to produce training samples called maps. In mobile networks, SOM can be used to find similar behavior cells located close to each other.
2	K-means clustering [73]	A combination of SOM and K-means clustering algorithm can be used to isolate the cell clusters with similar behaviors.
3.	Supervised binary classification [74]	This was used for identification of telecommunications fraud cases.
4.	Agglomerative hierarchical clustering [75]	Used in comparison with [71], where dendrograms were used to display the results.
5.	Naïve Bayesian classifier [76]	It has been used for automatic diagnosis for RANs which utilizes KPI information for diagnosis in GSM/GPRS networks.

Table 3.3 Some existing work done for detection of attacks

S.no.	Work done	Description
1.	Crowdroid [77]	Cloud-based machine learning framework for malicious application detection.
2.	MADAM [78]	This framework utilizes machine learning to classify applications. It extracted 13 features at the user and kernel level.
3.	DREBIN [79]	Method of detection of Android malicious applications on smart phones. It learns the difference between the benign and malicious applications by extracting the set of features and the disassembled code of applications.
4.	Andromaly [80]	Feedforward neural networks were used for training the model that builds distinguishable patterns between benign and malicious application. Recurrent neural networks were used to train the model with system calls of the benign applications execution behavior.
5.	Marvin [81]	This method combined the static and dynamic analysis and applied machine learning techniques to assess the risk of unknown applications.

3.4.3 Deep Learning for Detection of Security Threats

While mobile edge computing is an interesting computing paradigm for dynamic systems, new types of vulnerabilities and attacks (especially those that come from the edge network itself) are also one of the important challenges to address. Therefore, detecting attacks and associating them based on their location information are also important. Machine learning algorithms have been used to detect such attacks, and there has been a significant work in this direction. Table 3.3 gives an insight into the existing work in the detection of attacks.

3.5 Conclusion and Future Directions

The ubiquity of mobile phones and the associated big data consist of large and complex datasets generated by the wide variety of applications and other sensor data involved. With the implementation of MEC near the mobile devices can help the data analytics extracting the meaningful information from the raw data on the mobile edge node which could be then sent to the cloud servers for further processing. This will reduce the computational load on the cloud server for storage, data sync, and sharing. Deployment of MEC can solve many problems involving bandwidth, battery life, and storage to the resource-constrained mobile devices. In this chapter, we had a brief discussion on how the machine learning methods can help tailor the needs of the mobile edge analytics to overcome various challenges of mobile sensing, detection of the base station routers, and also the detection of attacks for secure and reliable solutions. For the future work, it would be interesting to investigate other deep learning algorithms for big data processing in MEC environment and apply these algorithms on different datasets to examine their effectiveness. Comparison of performance of different canonical approaches can be done, and potential applications of more sophisticated deep learning techniques can be explored. MEC still being in early stages of research and development have many challenges and, at the same time, give a variety of innovative opportunities in the future.

References

1. A. AL-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun. Sur. Tutorials **17**, 2347–2376 (2015)
2. M. Peng, K. Zhang, Recent advances in fog radio access networks: performance analysis and radio resource allocation. IEEE Access J. **4**, 5003–5009 (2016)
3. CISCO, The internet of things how the next evolution of the Internet is changing everything, white paper (2017)
4. C. Systems, Cisco visual networking index: global mobile data traffic forecast update, 2015–2020 white paper. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html
5. M. Chiang, T. Zhang, Fog and IoT: an overview of research opportunities. IEEE Internet Things J. **3**, 854–864 (2016)
6. T.X. Tran, A. Hajisami, P. Pandey, D. Pompili, Collaborative mobile edge computing in 5G networks: new paradigms, scenarios and challenges. IEEE Commun. Mag, Special Issue on Fog Computing and Networking, 1–7 (2017)
7. Mobile C, C-RAN: the road towards green RAN. White paper (2011)
8. Y.C. Hu, M. Patel, D. Sabella, N. Sprecher, V. Young, Mobile edge computing a key technology towards 5G. ETSI White Paper, vol. 11 (2015)
9. Intel and Siemens Networks, Increasing mobile operators' value proposition with edge computing. Technical Brief (2013)
10. 5G Amricas, Understanding information centric networking and mobile edge computing (2016). http://www.5gamericas.org/files/3414/8173/2353/Understanding_Information_Centric_Networking_and_Mobile_Edge_Computing.pdf

11. ITU, Work Plan, timeline, process and deliverables for the future development of IMT (2015)
12. ITU, High level 5G architecture. Network softwarization, gaps (2013). http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx
13. ETSI, 5G – the 5G infrastructure public private partnership: the next generation of communication networks and services (2015)
14. ETSI, Mobile edge computing – introductory technical whitepaper (2014)
15. Intel, Real-world impact of mobile edge computing. White paper (2016)
16. S. Yi, C. Li, Q. Li, A survey of fog computing: concepts, applications and issues, in *Proceedings of the 2015 Workshop on Mobile Big Data*, ACM, New York, 2015. pp. 37–42
17. Y. Jararweh, A. Doulat, A. Darabseh, M. Alsmirat, M. Al-Ayyoub, E. Benkhelifa, Sdmec: software defined system for mobile edge computing, in *IEEE International Conference on Cloud Engineering Workshop (IC2EW)*, April 2016, pp. 88–93
18. M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, A. Neal, Mobile edge computing introductory technical white paper. White paper. Mobile Edge Computing (MEC) industry initiative (2014)
19. C.K. Park, Performance for radio access network in mobile backhaul network. J. Inst. Internet. Broadcast. Commun. **12**(6), 297–302 (2012)
20. A.H. Khan, M.A. Qadeer, J.A. Ansari, S.Waheed, 4G as a next generation wireless network, in *Future Computer and Communication ICFCC 2009*, 2009, pp. 334–338
21. CommVerge, Radio Access Network (RAN) Optimization (2016). Available: http://www.commverge.com/Solutions/SubscribersServicesManagement/RANOptimization/tabid/174/Default.aspx
22. J. Wu, Z. Zhang, Y. Hong, Y. Wen, Cloud radio access network (C-RAN): a primer. IEEE Netw. **29**(1), 35–41 (2015)
23. Y. Liu, M.J. Lee, Y. Zheng, Adaptive multi-resource allocation for cloudlet-based mobile cloud computing system. IEEE Trans. Mob. Comput. **15**(10), 2398–2410 (2016)
24. F. Akhbar, V. Chang, Y. Yao, V.M. Munoz, Outlook on moving of computing services towards the data sources. Int. J. Inf. Manag. **36**, 645–652 (2016)
25. C.S. Li, F. Darema, V. Chang, Distributed behavior model orchestration in cognitive internet of things solution. Enterp. Inf. Syst. (2017). https://doi.org/10.1080/17517575.2017.1355984
26. M. Mital, V. Chang, P. Choudhary, A. Pani, Z. Sun, Adoption of cloud based internet of things in India: a multiple theory perspective. Int. J. Inf. Manag. (2016). https://doi.org/10.1016/j.ijinfomgt.2016.02.011
27. International Data Corporation, IDC Reveals Worldwide Internet of Things Predictions for 2015. Press Release (2015)
28. K. Hong, D. Lillethum, U. Ramachandran, B. Ottenwalder, B. Koldehofe, Mobile fog: a programming model for large-scale applications on the internet of things, in *Proceedings of the Second ACM SIGCOMM Workshop on Mobile Cloud Computing MCC'13*, New York, NY, USA, 2013, pp. 15–20. [Online]. Available: http://doi.acm.org/10.1145/2491266.2491270
29. International I, Computer technology increasingly aids traffic management (2009). http://www.itsinternational.com/categories/detection-monitoring-machine-vision/features/
30. A.V. Dastjerdi, H. Gupta, R.N. Calheiros, S.K. Ghosh, R. Buyya, Fog computing: principles, architectures and applications. arXiv preprint arXiv:1601.02752 (2016)
31. Y. Cao, S. Chen, P. Hou, D. Brown, Fast: a fog computing assisted distributed analytics system to monitor fall for stroke mitigation, in *IEEE International Conference on Network, Architecture and Storage (NAS)*, 2015, pp. 2–11
32. A.S. Go, D. Mozaffarian, V.L. Roger, E.J. Benjamin, J.D. Berry, M.J. Blaha, S. Dai, E.S. Ford, C.S. Fox, S. Franco, et al., Heart disease and stroke statistics – 2014 update. Circulation **129**(3) (2014)
33. P.A. Heidenreich, J.G. Trogdon, O.A. Khavjou, J. Butler, K. Dracup, M.D. Ezekowitz, E.A. Finkelstein, Y. Hong, S.C. Johnston, A. Khera, et al., Forecasting the future of cardiovascular disease in the United States a policy statement from the American Heart Association. Circulation **123**(8), 933–944 (2011)

34. V. Stantchev, A. Barnawi, S. Ghulam, J. Schubert, G. Tamm, Smart items, fog and cloud computing as enablers of servitization in healthcare. Sens. Transducers **185**(2), 121 (2015)
35. S. Datta, C. Bonnet, Haerri, Fog computing architecture to enable consumer centric internet of things services, in *IEEE International Symposium on Consumer Electronics (ISCE)*, pp. 1–2
36. W. Li, Y. Zhao, S. Lu, D. Chen, Mechanisms and challenges on mobility- augmented service provisioning for mobile cloud computing. IEEE Commun. Mag. **53**(2), 978–995 (2015)
37. H. Hromic, D. Le Phuoc, M. Seerano, A. Antonic, I.P. Zarko, C. Hayes, S. Decker, Real time analysis of sensor data for the internet of things by means of clustering and event processing, in *Proceedings of the IEEE International Conference on Communications*, 2015, pp. 685–691
38. http://aws.amazon.com
39. https://azure.microsoft.com/en-gb/
40. https://appengine.google.com/
41. B. Han, V. Gopalakrishnan, L. Ji, S. Lee, Network function virtualization: challenges and opportunities for innovations. IEEE Commun. Mag. **53**(2), 90–97 (2015)
42. A. Manzalini, R. Minerva, F. Callegati, W. Cerroni, A. Campi, Clouds of virtual machines on edge networks. IEEE Commun. Mag. **51**(2), 148–155 (2013)
43. C. Clark, K. Fraser, S. Hand, G. Hansen, E. Jul, C. Limpach, I. Pratt, A. Warfield, Live migration of virtual machines, in *Proceedings of the USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Berkley, USA, 2005
44. M. Habib ur Rehman, P.P. Jayaraman, S.R. Malik, M. Medhat Gaber, RedEdge: a novel architecture for big data processing in mobile edge computing environments. J. Sens. Actuator Netw. **6**, 17 (2017)
45. C.K.S. Leung, R.K. MacKinnon, F. Jiang, Reducing the search space for big data mining for interesting patterns form uncertain data, in *IEEE International Conference on Big Data*, Washington, DC, USA, 2014, pp. 315–322
46. A. Stateczny, M. Wlodarczyk-Sielicka, Self-organizing artificial neural networks into hydrographic big data reduction process, in *Rough Sets and Intelligent Systems Paradigms*, Spirnger, Berlin Germany, 2014, pp. 335–342
47. A. Ragyanszki, K.Z. Gerlei, A. Suranyi, A. Kelemen, S.J.K. Jensen, I.G. Csizmadia, B. Viskolcz, Big data reduction by fitting mathematical functions: a search for appropriate functions to fit Ramachandran surfaces. Chem. Phys. Lett. **625**, 91–97 (2015)
48. M. Schuster, Speech recognition for mobile devices at Google, in *Pacific Rim International Conference on Artificial Intelligence*, Springer, 2010, pp. 8–10
49. M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L.D. Jackel, M. Monfort, U. Muller, J. Zhang, et al, End to end learning for self- driving cars. arXiv preprint arXiv:1604.07316 (2016)
50. Google, Google cloud vision api (2017). https://cloud.google.com/vision. Accessed 11 Nov 2017
51. Movidius, Movidius neural compute stick (2017). https://developer.movidius.com/. Accessed 11 Nov 2017
52. M. Chen, U. Challita, W. Saad, C. Yin, M. Debbah, Machine learning for wireless networks with artificial intelligence: a tutorial on neural networks. arXiv preprint arXiv:1710.02913 (2017)
53. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016)
54. K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on ImageNet classification, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1026–1034
55. A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Proces. Syst. (NIPS) **25**, 1097–1105 (2012). Curran Associates Inc
56. K. Katevas, I. Leontiadis, M. Pielot, J. Serra, Practical processing of mobile sensor data for continual deep learning predictions, in *Proceedings of the 1st International Workshop on Deep Learning for mobile systems and applications*, NY, USA, 2017
57. S. Yao, S. Hu, Y. Zhao, A. Zhang, T. Abdelzaher, DeepSense: a unified deep learning framework for time-series mobile sensing data processing, in *International World Wide Web Conference Committee (IW3C2). Creative Commons CC BY,* 2017

58. A. Graves, Generating sequences with recurrent neural networks. ArXiv:1308.0850 (2013)
59. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. arXiv:1512.03385 (2015)
60. J. Donahue, A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, T. Darell, Long-term recurrent convolutional networks for visual recognition and description, in *CVPR*, 2015
61. K. Greff, R.K. Srivastava, J. Koutnik, B.R. Steunebrink, J. Schmidhuber, LSTM: a search space odyssey. arXiv:1503.04069 (2015)
62. N. Srivastava, R.R. Salakhutdinov, Multilodal learning with deep Boltzmann machines, in *NIPS*, 2012
63. S. Bhattacharya, N.D. Lane, From smart to deep: Robust activity recognition on smartwatches using deep learning, in *PerCom Workshops*, 2016
64. V. Radu, N.D. Lane, S. Bhattacharya, C. Mascolo, M.K. Marina, F. Kawsar, Towards multimodal deep learning for activity recognition on mobile devices, in *UbiComp: Adjunct,* 2016
65. D. Figo, P.C. Diniz, D.R. Ferreira, J.M. Cardoso, Preprocessing techniques for context recognition from accelerometer data. Pers. Ubiquit. Comput **14**, 645–662 (2010)
66. C.-Y. Li, C.-H. Yen, K.C. Wang, C.-W. You, S.-Y. Lau, C.C.-H. Chen, P. Huang, H.-H. Chu, Bioscope: an extensible bandage system for facilitating data collection in nursing assessments, in *UbiComp*, 2014
67. E. Miluzzo, A. Varshavsky, S. Balakrishnan, R.R. Choudhury, Tapprints: your finger taps have fingerprints, in *MobiSys*, 2012
68. W.D. Fisher, Machine learning for the automatic detection of anomalous events, PhD dissertation, Colorado School of Mines, 2017, pp. 173
69. J. Laiho, M. Kylvaja, A. Hoglund, Utilization of advanced analysis methods in UMTS networks, in *IEEE Vehicular Technology Conference*, Birmingham, USA, 2002, pp. 726730
70. J. Laiho, K. Raivio, P. Lehtimaki, K. Hatonen, O. Simula, Advanced analysis methods for 3G cellular networks. IEEE Trans. Wirel. Commun. **4**(3), 930942 (2005)
71. A.J. Hoglund, K. Hatonen, A.S. Sorvari, A computer host based user anomaly detection system using the self-organizing map, in *IEEE-INNS-ENNS International Joint Conference on Neural Networks, vol. 5*, Italy, 2005, pp. 411416
72. P. Lehtimaki, K. Raivio, A SOM based approach for visualization of GSM network performance data, in *International Symposium on Intelligent Data Analysis*, Madrin, Spain, 2005
73. G. Chen, S.A. Jaradat, N. Banerjee, T.S. Tanaka, M.S. Ko, M.Q. Zhang, Evaluation and comparison of clustering algorithms in analyzing ES cell gene expression data. Stat. Sin. **12**(1), 241–262 (2002)
74. C.S. Hilas, P.A. Mastorocostas, An application of supervised and unsupervised learning approaches to telecommunications fraud detection. Knowl. Based Syst. **21**(7), 721–726 (2008)
75. J.H. Ward Jr., Hierarchical grouping to optimize an objective function. J. Am. Assoc. **58**(301), 236–244 (1963)
76. J. You, Research of wireless network fault diagnosis based on Bayesian networks, in *Second International Symposium on Knowledge Acquisition and Modeling, vol. 3*, 2009, pp. 59–64
77. I. Burguera, U. Zurutuza, S. Nadjm-Tehrani, Crowdroid: behavior-based malware detection system for android, in *ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, ACM, 2011, pp. 15–26
78. G. Dini, F. Martinelli, A. Saracino, D. Sgandurra, Madam: a multi-level anomaly detector for android malware, in *Computer Network Security*, 2012, pp. 240–253
79. D. Arp, M. Spreitzenbarth, M. Hubner, H. Gascon, K. Rieck, C. Siemens, Drebin: effective and explainable detection of android malware in your pocket, in *NDSS*, 2014
80. F. Al-Turjman, S. Alturjman, Confidential smart-sensing framework in the IoT era. J Supercomput **74**(10), 5187–5198 (2018)
81. M. Lindorfer, M. Neugschwandtner, C. Platzer, Marvin: efficient and comprehensive mobile app classification through static and dynamic analysis, in *39th IEEE Annual Computer Software and Applications Conference (COMPSAC), vol. 2*, IEEE, 2015, pp. 422–433

Chapter 4
Taxonomy of Edge Computing: Challenges, Opportunities, and Data Reduction Methods

Kusumlata Jain and Smaranika Mohapatra

4.1 Introduction

In recent years, cloud computing has made significant advances and is being more commonly used because of its high value efficiency and flexibility resulting from consolidation. Cloud computing has led to the speedy development of cellular internet and internet of things (IoT) packages for demanding situations using sophisticated software and services. IoT is experiencing explosive growth and is virtually transforming all aspects of modern life [1]. Cisco has predicted that 50 billion devices will be linked to the IoT by 2020 [2]. As these new technologies enter the everyday lives of consumers, new challenges arise that cannot be safely addressed through a centralized cloud computing structure, including stringent latency, capacity constraints, useful resource-confined devices, uninterrupted services with intermittent connectivity, improved security [3], and latency needs. Because the records increase each day on the network, the performance of the machines continues to be affected. This chapter discusses basic information about cloud technology and IoT, associated challenges, and statistics reduction techniques for edge computing to reduce data and latency time.

K. Jain (✉) · S. Mohapatra
Department of Computer Science & Engineering, Maharishi Arvind Institute of Engineering & Technology, Jaipur, Rajasthan, India

Rajasthan Technical University, Kota, Rajasthan, India

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_4

4.2 Cloud Computing

The term "cloud" refers to a network or internet that is present at a remote location. The cloud can provide changes over public and private networks. Cloud computing is about controlling/moving, configuring, and using hardware and software remotely. It offers online data storage and the basic functionality needed for a business or society to operate. However, certain accommodations and models may be needed locally for cloud computing to be feasible and accessible to end users. Two categories of models for cloud computing are the deployment models and service models.

The deployment models describe usage access to the cloud. The cloud may use one of four types of deployment models: public, private, hybrid, or community:

- *Public cloud:* A public cloud sanctions systems and changes places to easily accommodate the general public. A public cloud may be less secure than other types of clouds because of its accessibility and openness.
- *Private cloud:* In a private cloud, systems and services are easy to access, use, and understand within an organization. A private cloud is more secure than other clouds because of its private nature.
- *Hybrid cloud:* A hybrid cloud has characteristics of both private and public clouds. The analytical process is performed through a private cloud, whereas non-analytical work is performed through a public cloud.
- *Community cloud:* A community cloud is deployed by a group of organizations or industries.

The service models describe the eight types of services offered by a cloud:

- *Everything-as-a-Service (EaaS):* This type of service is given to all users of the software and hardware to control business processes, including interactions between users. The user only needs to have access to the internet.
- *Infrastructure-as-a-Service (IaaS):* In this type, users can avail a service or computing platform by using a virtual platform that is connected to the network. The service is automatically adjusted according to the use.
- *Platform-as-a-Service (PaaS):* In this service, the user is given a computing platform along with operating software and other necessary programs.
- *Software-as-a-Service (SaaS):* Also known as "software on demand," this service deploys software on remote servers but also uses the internet.
- *Hardware-as-a-Service (HaaS):* This service uses uncovered hardware on which a user can deploy its own software, similar to IaaS but deviating a bit.
- *Workspace-as-a-Service (WaaS):* This service is implemented by organizations or companies that use cloud computing to set up the necessary workspace for their employees.
- *Data-as-a-Service (DaaS):* This service is used to store large amounts of data and information in a storage space.

- *Security-as-a-Service (SaaS):* This service enables users to safely implement and use Web technologies, providing security in communication media and local networks.

4.3 Internet of Things

IoT is an advanced automation and analytical system that has exploited various aspects of technology, including recent advances in software enhancements, falling hardware prices, and modern attitudes toward technology. IoT combines sensing, communication, astronomically immense data, and artificial intelligence to distribute consummate systems for a product or accommodation. This collection of new and advanced technologies has made major behavioral, political, and socioeconomic changes in the lives of techno-savvy consumers. The most important feature of IoT is its ability to make things “smart”; in this way, it enhances every aspect of life.

The “thing” in IoT can be any contrivance that is artificially smart enough to communicate with a network without manual intervention to facilitate the interaction. The communication can be with internal states of the “thing” or the external environment, which in turn assists with the decision-making process. IoT allows a “thing” to be smart enough to make a decision for you; for example, your refrigerator and cabinets can become smart enough to detect when milk and your favorite cereal are running low, then place an order with your preferred grocer. As such, IoT introduces a new paradigm for active content, product, or accommodation engagements. IoT devices have become smaller, more affordable, and more powerful over time. IoT uses these purpose-built minuscule devices to work with precision, scalability, and diversity.

The advantages of IoT span across every area of lifestyle and business to improve the overall quality of life, including in communications, automation and control, data gathering, data monitoring, time saving, cost-effectiveness, and efficiency. Despite all its advantages, IoT also has security and privacy issues. The major disadvantage of IoT is that it requires immense data transmission to the cloud and thus increased security.

IoT is increasing exponentially and will potentially network billions or even trillions of devices. Ericsson Inc. [4] has estimated that the data produced by people, machines, and things will reach 500 zettabytes by 2019; however, Cisco Global Cloud Index has estimated that global data centre IP traffic will only reach 10.4 zettabytes by that time [5]. Most IoT devices will be located at the edge of the internet and could provide new applications, changing many aspects of both traditional industrial productions and our everyday living. Some of these devices are already on the market, including Apple watches, Oculus Rift helmets [6], Google Nest [7], Fitbit sports trackers, and Google Glasses. By 2019, 45% of IoT-created data will be stored, processed, analyzed, and acted upon close to (or at the edge of) the network [8] but eventually transmitted to the cloud.

4.4 Machine-To-Machine Communication

In machine-to-machine communication, multiple devices can connect and exchange information directly with each other, without any intermediary hardware assistance [9]. These devices can connect with each other over different types of networks, including but not limited to the internet or IP networks. These device-to-device networks allow devices to exchange information in hybrid communication protocols, which combine device-to-device and other communication protocols to achieve quality-of-service (QoS) requirements. This model is commonly used in applications such as smart home systems and automatic controls in electrical systems, which communicate with each other by sending small data packets with relatively low data rate requirements. Some typical IoT devices of this type are smart door locks, smart switches, and smart lights, which also typically only exchange small data packets.

From the user's perspective, a problem with machine-to-machine communications is a lack of compatibility, because different devices from different manufacturers use different protocols. Taking smart home devices as an example, Z-Wave protocol devices cannot communicate with ZigBee protocol devices [10]. These compatibility issues limit a user's options and experience.

4.5 Edge Computing

As IoT grows, the data in the network is also increasing. To reduce these data, the technology of edge computing may be used. Edge computing allows devices to perform computations at the edge of the network, on downstream data on behalf of cloud services and on upstream data on behalf of IoT services. Here, "edge" refers to any computing and network resources that are between the data sources and cloud data centers. For example, a smart phone is the edge between an individual and the cloud, a gateway in a smart home is the edge between the house and the cloud, and a micro data center and a cloudlet are the edge between a mobile device and the cloud. Rationally, edge computing is performed at proximity to the data source. Figure 4.1 illustrates the concept of edge computing.

In the edge computing paradigm, the "things" are data consumers and data producers. The edge can perform computing offloading, data storage, caching and processing, request distributions, and distribution accommodations from the cloud to the user. The edge has QoS requirements such as reliability, security, and privacy firewalls.

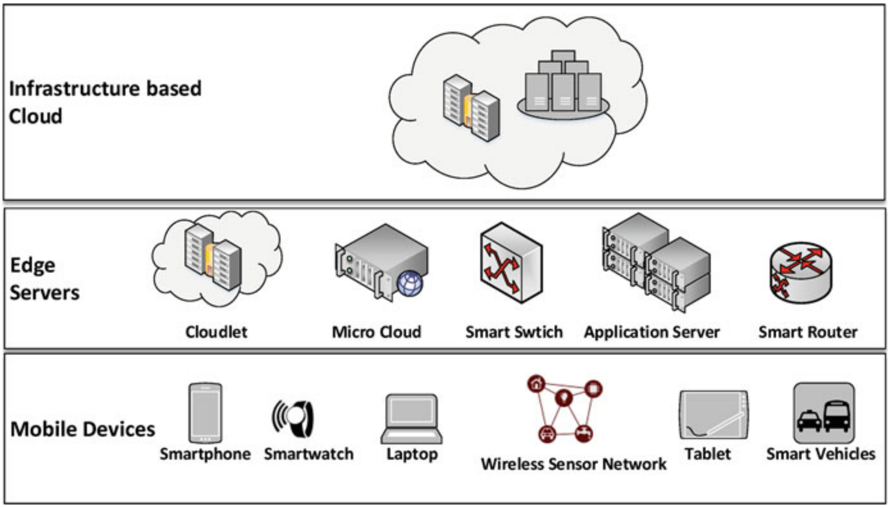


Fig. 4.1 Edge computing paradigm

4.5.1 Why Edge Computing

IoT applications generate enormous amounts of data by IoT sensors. The big data are subsequently analyzed to determine reactions to events or to extract analytics or statistics. However, sending all of these data to the cloud would require prohibitively high network bandwidth. Current research efforts are investigating how to better exploit capabilities at the edge of the network to support the IoT and its needs [11]. In edge computing, the massive data generated by different kinds of IoT devices can be processed at the network edge instead of transmitting the data to a centralized cloud infrastructure due to bandwidth and energy consumption concerns. Edge computing can provide services with faster response and better quality than cloud computing. Edge computing is suitable for integration with IoT to provide efficient and secure services for many end users. Edge computing-based architecture also should be considered for future IoT infrastructure.

Push from Cloud Services Putting all computing tasks on the cloud is an efficient way to process data because the cloud's computing power outclasses the capability of the things at the edge. However, compared with the fast-developing speed of data processing, the bandwidth of the network has come to a standstill. With the growing quantity of data generated at the edge, the speed of data transportation is becoming a bottleneck in the cloud-based computing paradigm. For example, about 5 gigabytes of data are generated by a Boeing 787 every second [12], but the bandwidth between the airplane and either a satellite or the base station on the ground is not large enough for data transmission. Consider an autonomous vehicle as another example: 1 gigabyte of data is generated by the car every second, and it requires real-time processing for the vehicle to make correct decisions [13]. If all the data needed to be sent to the cloud for processing, the response time would be too long. Furthermore, it would be challenging for current network bandwidth and reliability to support many vehicles in one area. In this case, the data needs to be processed at the edge for shorter response times, more efficient processing, and reduced network pressure.

Pull from IoT Many kinds of electrical devices will become part of the IoT. These devices will play the roles of both data producers and data consumers, such as air quality sensors, LED bars, streetlights, and internet-connected microwave ovens. It is safe to infer that the number of things at the edge of the network will increase to the billions in a few years. The raw data produced by these devices will be enormous, with conventional cloud computing being too inefficient to handle the data. Thus, most of the data produced by IoT will never be transmitted to the cloud. Instead, the data will be consumed at the edge of the network.

In conventional cloud computing, data producers generate raw data and transfer it to the cloud, and data consumers a send request for consuming data to the cloud. However, this structure is not sufficent for IoT. First, the data quantity at the edge is too large, which will lead to unnecessary bandwidth and computing resource usage. Second, privacy protection requirements will pose an obstacle for cloud computing in IoT. Lastly, most of the end nodes in IoT are energy-constrained things, whereas the wireless communication module is usually very energy hungry. Therefore, offloading some computing tasks to the edge could be more energy efficient.

Change From Data Consumer to Producer In the cloud computing paradigm, the end devices at the edge usually as as data consumers, such as when you watch a YouTube video on your smart phone. However, people now are also producing data from their mobile devices. This change from data consumer to data producer/consumer requires more function placement at the edge, such as when users take photos or record video.

4.5.2 Technologies to Implement Edge Computing

This section discusses the basic techniques to implement edge computing. As edge computing pushes the data storage and process near the edge of the network, small data centers can be provided near the edge device, with or near the cloud, in techniques known as cloudlets, mobile edge computing (MEC), and fog computing.

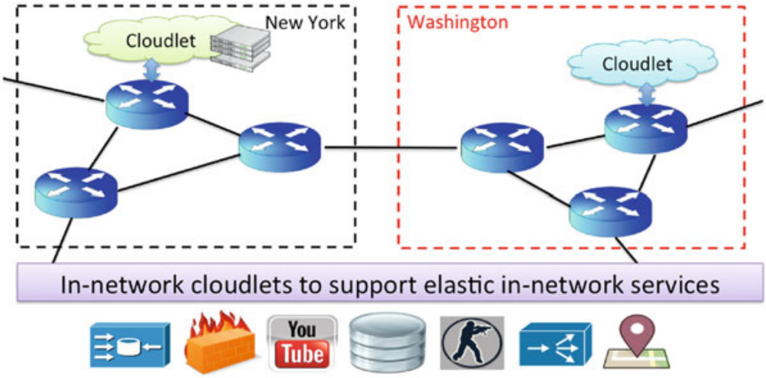


Fig. 4.2 Cloudlet architecture

4.5.2.1 Cloudlet

A cloudlet may address the critical challenge of end-to-end responsiveness between a mobile device and the associated cloud. A cloudlet is a mobility-enhanced small-scale cloud data center that is located at the edge of the internet. It is a trusted, resource-rich system or cluster of systems that are well connected to each other and the internet, which is available for use by nearby mobile devices [14]. As shown in Fig. 4.2, a cloudlet is the middle tier in a three-tier hierarchy architecture to achieve good response times.

The aim of cloudlet is to provide lower latency by supporting resource-intensive and interactive mobile applications through powerful computing resources for mobile devices. Mobile devices use a nearby cloudlet for computing toward the cloud. It is important here to differentiate a cloudlet from a cloud. First, a cloudlet is much more agile in its provisioning because associated edge devices are highly dynamic and mobile. Second, a good handoff is needed to support the seamless mobility of edge devices. Third, to connect or communicate with the cloud, a mobile node first connects to a nearby cloudlet and then the cloud in order to complete its communication. Cloudlets reduce data on the network and the response latency to mobile nodes.

4.5.2.2 Mobile Edge Computing

The concept of MEC was defined by the European Telecommunications Standards Institute (ETSI) as a new technology that "provides an IT service environment and cloud-computing capabilities at the edge of the mobile network, within the Radio Access Network (RAN) and in close proximity to mobile subscribers" [15]. In a published white paper on MEC, ETSI considers MEC to be a key emerging

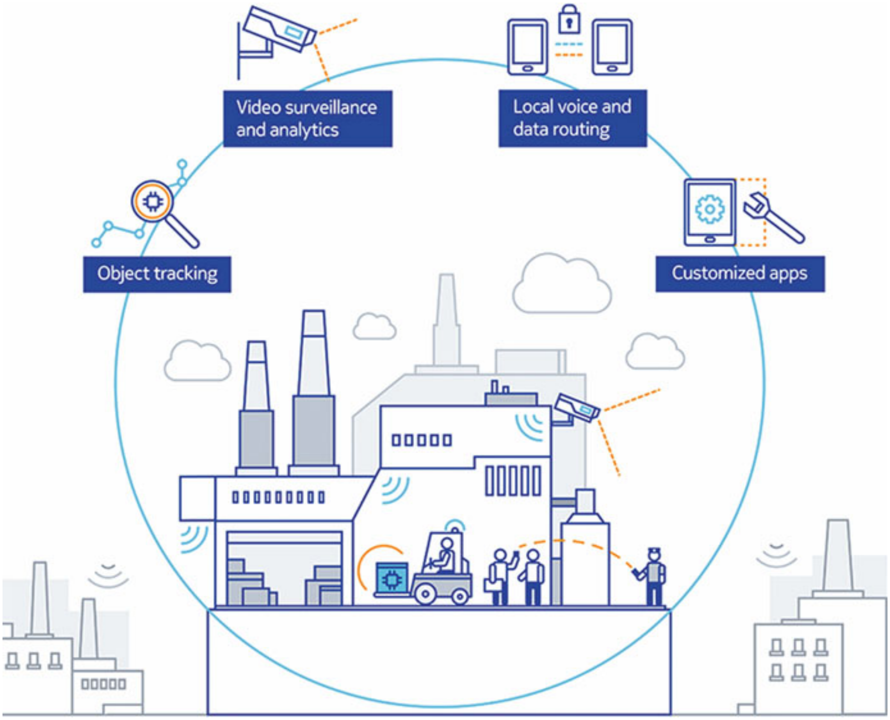


Fig. 4.3 Mobile edge computing

technology and an important component of future generation networks [16]. Figure 4.3, illustrates the concept of MEC.

In December 2014, ETSI established an Industry Specification Group (ISG) on MEC to develop a standardized, open environment that will allow efficient and seamless integration of third-party applications across multi-vendor platforms. In January 2017, the MEC ISG released six specifications—one of which provides a glossary of terms related to the conceptual, architectural, and functional elements of MEC [17]. The purpose of this specification is to enable the consistent use of terminology within ETSI MEC specifications and, beyond the ISG, more widely in the industry. Other specifications describe the technical requirements enabling interoperability and deployment, example use cases, and their technical benefits [18] and provide a framework and reference architecture to enable mobile edge applications to run efficiently and seamlessly in a mobile network [19]. The fourth specification in MEC ISG introduces a number of service scenarios that would benefit from the MEC technology [20]. The proof-of-concept (PoC) framework specification defines a framework to coordinate and promote multi-vendor PoC projects illustrating key aspects of MEC technology [21]. PoC projects are essential to demonstrate MEC as feasible and valuable, to validate the specifications that are being developed, to demonstrate use cases, and ultimately to help develop a diverse and open MEC ecosystem. The last specification describes various metrics that can potentially be improved by deploying a service on a MEC platform, such as latency, energy efficiency, network throughput, system resource footprint, and quality [22]. Furthermore, the last specification also describes the best practices for measuring such performance metrics.

ETSI announced six different MEC PoCs in September 2016, which were accepted at the MEC World Congress in Munich. These PoCs will help to strengthen the strategic planning and decision-making of organizations and identify which MEC solutions may be viable in the network. Due to its advanced characteristics, such as low latency, proximity, high bandwidth, and real-time insight into radio network information and location awareness, MEC enables many new kinds of applications and services for multiple sectors, such as consumer, enterprise, and healthcare. MEC also seems to be a promising solution for handling video streaming services in the context of smart cities [23].

4.5.2.3 Fog Computing

The new technology of fog computing is similar to edge computing but it works with the cloud. The OpenFog Consortium was founded to drive industry and academic leadership in fog computing architecture, testbed development, and a variety of interoperability and composability deliverables that seamlessly leverage cloud and edge architectures to enable end-to-end IoT scenarios [24]. OpenFog Consortium published a white paper on fog computing in February 2016, which outlined its approach to an open fog computing architecture (OpenFog architecture) [25]. The OpenFog Consortium defined fog computing as "a system-level horizontal architecture that distributes resources and services of computing, storage, control and networking anywhere along the continuum from Cloud to Things (https://www.openfogconsortium.org)." Fog computing differs from edge computing by providing tools for distributing, orchestrating, managing, and securing resources and services across networks and between devices that reside at the edge. Edge architecture places servers, applications, and small clouds at the edge. Fog computing jointly works with the cloud, whereas edge is defined by exclusion of the cloud. Figure 4.4 illustrates the concept of fog computing. In Figs. 4.3 and 4.4, fog computing and edge computing are shown as the middle layer; however, edge computing is toward the edge device, whereas the middle layer works with the cloud in fog computing.

Fog computing standardization is mainly administered by the OpenFog Consortium, whose objective is to influence standard bodies to create standards so that IoT systems at the edge can inter-operate securely with other edge and cloud services in a friction-free environment. The OpenFog Consortium has set up six working groups: an architecture working group, communications working group, manageability working group, security working group, software infrastructure working group, and testbed working group. These working groups evaluate, classify,

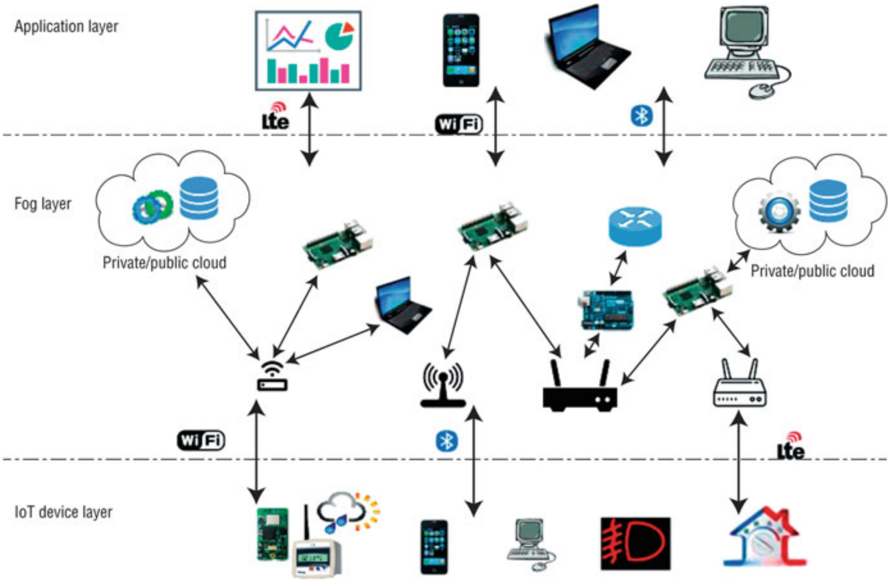


Fig. 4.4 Fog computing

and recommend standards, practices, and technologies that are appropriate for OpenFog architecture to address corresponding challenges.

4.6 Data Reduction Techniques

The expected huge increase in the number of IoT data sources, such as sensors, embedded systems, and personal devices, provides good reason to implement network-edge computing—that is, data pre-processing, local storage, and filtering close to the data source. Specifically, data reduction at the network edge that is on an IoT gateway device or a mini-server deployed locally (a cloudlet) or on an IoT area network can prevent I/O bottlenecks, as well as dramatically reduce storage, bandwidth, and energy costs. Many solutions have been provided, but these solutions have two main obstacles in the implementation. Firstly, the most efficient algorithms for data reduction in time series (which is one of the prevailing types of data in IoT) were developed to work posteriori on big datasets; thus, they cannot make decisions per incoming data item. Secondly, the state of the art lacks systems that can apply any of the possible data reduction methods without adding significant delays or heavyweight re-configurations.

There are many data reduction techniques used in edge computing. Based on when data are sensed, two main categories of approaches can be used for IoT data reduction at the network edge: (1) time series compression approaches or (2) event- or policy-based engines. Time series compression and aggregation techniques are based on sampling, summarization, approximation, or a similar approach. For example, Fast Time Sequence Indexing for Arbitrary LpNorms is based on piecewise approximation and replacement of every n points with their arithmetic means. Another method is based on selecting only the "perceptually important points" (PIPs) of the time series, which measures importance based on the distance from other PIPs and on local minima and maxima, respectively. However, such solutions are not designed to act upon streams (i.e., per incoming data item), but rather on already collected data sets (posteriori). Furthermore, they are not complemented by data quality control mechanisms; thus, they are often avoided because of the information loss that selective forwarding or data filtering inherently implies. Finally, to the best of our knowledge, they have not been sufficiently applied or evaluated on modern types of IoT time series, such as sensor data from traffic sensors, smart meters, industrial automation, or even smartphones.

Data handling solutions that use event- or policy-based engines have often (implicitly) a similar final goal as time series compression approaches—namely data reduction, although they usually rely on domain-specific rules rather than on time series characteristics. For example, Shen et al. [26] described how to use IP multimedia subsystems policies in order to perform (among others) a kind of data aggregation on network-edge (gateway) devices in machine-to-machine systems. Furthermore, Krikkit [27], an open-source solution initiated by Cisco, is in the process of specifying a data format and a mechanism for telling network-edge devices which data to forward and how. However, these approaches stop at designing policy engines and policy languages for data reduction, without developing or providing actual data reduction techniques. Thus, they practically require the network edge developer to implement the data reduction logic from scratch. A recent solution called FlexAnalytics [28] took things a step further by providing flexible placement of analytics for data reduction between the network edge and the backend system. This approach also includes experiments showing how different data reduction techniques (mainly compression and visualization-based data reduction) perform when applied at the network edge. However, it does not enable new data reduction techniques (or ease existing ones when they are applied at the network edge).

The three data reduction techniques described in the following sections can reduce data collection and transmission process times.

Data Sampling Data sampling either reduces the amount of sensory data to be further processed, paying special attention to extracting the most representative points of the data set, or chooses a subset of sensors that contribute sensory data among all the IoT sensors. Willett et al. [29] proposed an adaptive sampling approach, in which only few sensor nodes are first activated to sample, which leverages spatial correlations. Higher correlations (i.e., IoT sensors are closer to each other) require fewer active sensors to sample data. The data are then sent to a data fusion center, which derives valuable information regarding the environment and the phenomenon under observation, activating more nodes if needed. The data fusion center allocates network resources accordingly. By contrast, Jain and Chang [30] adapted the sampling rate of the sensors to the communication resources while minimizing the active sensors that stream data to a remote entity (i.e., server).

Compression Minimized datasets are easier to use in terms of processing and inside the network movement for large storage. Compression-based techniques are useful and appropriate for data reduction in edge computing. Compression can be applied to already-sensed data. However, Li et al. [31] proposed a compressed sensing framework, in which compressed sensing is applied prior to data acquisition. In an approach by Liu et al. [32], compressed sensing is applied to the data on-the-fly, under the assumption of opportunistic routing.

Data Aggregation Data aggregation consists of considering diverse sources of data and gathering them together to build an accurate representation of the phenomena under observation. These sources differ in that each of them is responsible for a specific sensing task. In IoT and wireless sensor networks, data aggregation techniques include operations such as solving the maximum, minimum, sum, average, median, and the count. Aggregation methods can be divided into centralized and distributed methods. Centralized methods require continuous communication among network entities. Hence, the communication overhead incurs additional costs; consequently, such methods are not suitable for sensor networks. Therefore, distributed methods such as clustering, multipath, and aggregate trees are commonly employed [33]. Bash et al. [34] performed data aggregation with the purpose of minimizing messaging costs among the sensor nodes. The node from which the data is aggregated is proportional to the given size of the network. Similarly, Lin et al. [35] divided the sensor network into non-overlapping regions, then sampled and aggregated data from each region. Moreover, Kimura and Latifi [36] summarized several methods of data compression, which consists of transmitting data as little as possible by aggregating. Such methods are often valid for specific lookup applications; hence, they are not applicable for applications that rely on fast- and ever-changing sensory data.

Most data aggregation methods are based on trees or other fixed data structures. For example, Huang et al. [37] proposed a spanning tree-based data aggregation mechanism for large-scale networks. Each leaf in the tree is responsible for sensing data from a given location. Data aggregation starts from the leaves and propagates to the root (i.e., data collection entity). Other data aggregation methods rely on no specific data structures [38]. Multi-task data sharing on the network leads to efficient use of the available bandwidth; however, it introduces more computation and communication costs. Such a technique aims at collecting as little data as possible to satisfy all tasks' needs.

IoT network architectures usually have a sensor-gateway router-cloud structure [39]. Indeed, such an architecture tailors the proposed mechanism for data reduction, with an aim to reduce the amount of transmitted data (and hence transmission power) at a lower time cost. Pratt and Fink [40] proposed a technique for fast compression of time series data and indexing of the compressed series. The compression mechanism identifies the important points of a series and discards the remaining ones. This technique has a good performance and can also be applied to resource-concerned applications. This solution proposed a framework for real-time data filtering over two tiers—namely, the gateway and the edge tier.

Other Approaches Other methods are being used to optimize data redundancy and duplicity. Some techniques use the topological structures of unorganized data and minimize the overall data using the network theory approaches:

1. *Network Theory*: The graph or networks play a basic role in reduction techniques of higher dimensional unorganized data into lower dimensional organized data [41]. However, the extraction of these structures is a difficult task due to heterogeneity and complex structures. Trovati [42] proposed a network theory approach that can abstract the properties of a network from the data. The topological networks are constructed by establishing and evaluating relationships (links) among different data points. The statistical node analysis of the networks is performed for optimization and immensely colossal data reduction [43]. The optimized networks are used in small networks, free scale, and any random network.
2. *Data Redundancy Elimination*: Data redundancy is a key issue for data analysis in immensely colossal data environments. Three main reasons for data redundancy are the integration of nodes, the expansion of datasets, and data replication. The integration of a single virtual machine (VM) brings around 97% more redundancy, and the magnification in immensely colossal datasets comes with 47% redundant data points [44]. Cluster deduplication may be a generalized massive information reduction scheme for disk-based cluster backup systems. The redundant information held on multiple disks and partition areas is a serious challenge for giant processing systems. The deduplication techniques salow systems to handle completely different information chunks (partitions) and exploitation hash functions with lower intra-node and inter-node communication overheads. Additionally, these strategies improve the storage potency by eliminating redundant information from multiple nodes.
3. *Data Pre-processing*: Data pre-processing techniques help to extract meta-data for further processing. Di Martino et al. [45] discussed the basic approaches for data pre-processing, which are generally dependent on semantic analysis or data structures. Some other techniques are also present, such as low memory pre-filters for streaming data, URL filtration, and map-reducing methods.
4. *Data Mining and Machine Learning*: Recently, many data mining and machine learning techniques have been used for data reduction. This concept is either implemented to reduce data instantly after the acquisition or customized to address some specific tasks. Jiang et al. [46] proposed a context-aware big data forwarding scheme using distributed wearable sensor devices. The proposed technique minimizes the communication and storage overhead in big data.

4.7 Some New Methods

Papageorgiou et al. [47] presented a new data reduction solution that automates the switching between different data handling algorithms at the network edge, including an analysis of adjusted data reduction methods. The authors also proposed three versions of a new algorithm that is capable of performing real-time reduction of incoming time series items based on the concept of perceptually important points.

Rehman et al. [48] presented and evaluated a novel big data processing architecture named RedEdge (i.e., data reduction on the edge) that incorporates a mechanism to facilitate the processing of big data streams near the source of the data. The RedEdge model leverages mobile IoT-termed mobile edge devices as primary data processing platforms. In the case of the unavailability of computational and battery power resources, it offloads data streams in nearer mobile edge devices or to the cloud. The RedEdge architecture and the related mechanism were evaluated within a real-world experiment setting involving 12 mobile users.

Kim et al. [49] took the approach of moving control of a low-power wide area network (LPWAN) to the edge cloud, which provides computing and storage environments at base stations. An LPWAN gateway is then integrated with the edge cloud, and LPWAN data are cached at the edge cloud in the gateway for LPWAN control. LPWAN controls, such as report interval control for data, transmission power control, and data aggregation, are applied to improve the efficiency of data transmission. Network control is performed by using cached data in the edge cloud. Compared with the existing LPWAN control approach, the proposed approach exhibits improved performance for IoT data transmission. The simulation results demonstrated the proposed approach’s efficiency.

4.8 Comparison of Various Data Reduction Techniques

A comparison of various data reduction techniques is provided in Table 4.1.

4.9 Challenges and Solutions in Edge Computing

Programmability Generally in a cloud environment, users program the code and deploy it in the cloud. The cloud provider is responsible for deciding where to perform the computing. The users are unaware of the computation. In edge computing, the computation is offloaded from the cloud and the nodes at the edge are heterogeneous, so the runtimes of these links differ. To address the programmability issue, a computing stream defined as a series of routines is applied to the data along with the propagation path. The computing stream can help the user to determine the function to be done and how the data is propagated after the computing is done at the edge.

Table 4.1 Comparison of Data Reduction Techniques

Methods	Description	Advantages	Disadvantages
Spatiotemporal [50]	Performs online clustering of streaming data by relating the similarities in time series. Performs temporal data compression on each of the network links to reduce the data.	Increased performance, including data quality and data reduction	Processing is required to be done on the edge and cloud computing at least once
gZIP [51]	Used for data reduction to increase the resource performance and efficiency for IP-datasets	Lightweight and has a simple file format, so there is low complexity in computational terms	Compresses one file at a time, so other techniques are used to increase the computation time
AST [44]	Used to compress digital signals by stretching and wrapping methods	Performs its operations using frequency	Only works with signals
RED encoding [52]	Manages massive and huge voluminous data	High compression rates are used	Performance degrades with high variance of signal
DQC [53]	Uses a visual data mining methodology	Can expose inner hidden structures and can signify larger dimensional data	Lacks performance and requires tools
TNs [54]	Uses sensor decomposition and approximation	Reduces feature spaces	High complexity

Naming The naming system is very important for programming, addressing, and identification, among others. A naming scheme is required to handle the mobility of things, dynamic network topology, and security issues in edge computing. Traditional naming systems, such as DNS, are not able to adequately handle dynamic edge computing. To overcome this issue, new naming mechanisms, such as named data network (NDN) [55] and MobilityFirst [56], can be deployed in edge computing. NDN provides a hierarchical structured name for the data network and is user friendly. MobilityFirst can separate the name from the network address to provide better mobility support and is efficient when applied to the edge service.

Data Abstraction With the use of IoT, a huge number of data generators are present in the data network. Human interaction in edge computing should be reduced; the edge links and nodes should use all the data and interact with the users. In this case, the data should be pre-processed at the gateway. Collecting the data will serve the application, and the process should be allowed to control the things to complete the various services. Combining data representation and operations as a layer of data abstraction serves as a public interface for all things connected with the edge operating system.

Service Management To ensure the reliability of a system, the following four features are required:

1. *Differentiation*: Multiple services are deployed at the edge network. These services have different priorities.
2. *Extensibility*: The things in IoT can be dynamic in nature. A newly owned thing should easily be added to the current service. In addition, when a thing is replaced, the previous service should be able to adopt it easily.
3. *Isolation*: If an application fails in a mobile operating system, the whole system crashes and needs to be rebooted. In edge computing, this can be tedious and complicated. This issue can be resolved by introducing a deployment and undeployment framework.
4. *Reliability*: The reason for a service failure is sometimes unknown. However, reliability is required to maintain the network topology of the whole system. Data sensing and communication also need to be reliable.

For the above issues, some protocols has been implemented for communication of IoT data [57].

Privacy and Security At the edge of the network, security and privacy are big issues and challenges. In some cases, intruders may know specifications of the data and may harm the privacy of the users. Some private information can be removed from data before processing.

4.10 Conclusion

Today, many devices have been pushed from the cloud to the edge of the network. When data are processed at the edge, shorter response times and greater reliability can result. Bandwidth also can be saved when a larger part of data is handled at the edge rather than uploaded to the cloud. The IoT and mobile devices have driven the trend of edge computing techniques from data consumer to data producer/consumer, where it is efficient to compute the data at the edge of the network. This chapter discussed the basics of edge computing, its various methodologies, and some data reduction techniques. Finally, we presented some of the challenges of edge computing and solutions to overcome them, including programmability, naming, data abstraction, privacy, and security.

References

1. A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutor **17**, 2347–2376 (2015)
2. CISCO, The Internet of Things How the Next Evolution of the Internet Is Changing Everything, White Pap. (2011). http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

3. M. Chiang, T. Zhang, Fog and IoT: an overview of research opportunities. IEEE Internet Things J. **3**, 854–864 (2016)
4. Ericsson Inc, Ceo to shareholders: 50 billion connections 2020, p. 1, (2010). Available at: http://www.ericsson.com/thecompany/press/releases/2010/04/1403231
5. Cisco global cloud index: Forecast and methodology, 2014–2019 white paper. (2014)
6. Oculus, Oculus Rift helmet: next generation virtual reality. (2016). Available at: https://www3.oculus.com/en-us/rift/. 2
7. Google, Nest IoT devices. (2016). Available at: https://nest.com/.3
8. D. Evans, The Internet of Things: how the next evolution of the Internet is changing everything. CISCO White Paper, vol. 1, pp. 1–11, 2011.10
9. F. Wortmann, K. Flüchter, Internet of Things. Bus. Inf. Syst. Eng. **57**(3), 221–224 (2015)
10. https://vtechworks.lib.vt.edu/bitstream/handle/10919/78767/Kalin_JH_T_2017.pdf?sequence=1
11. A. Brogi, S. Forti, QoS-aware deployment of IoT applications through the fog. IEEE Internet Things J. **4**, 1185–1192 (2017)
12. Boeing 787s to Create Half a Terabyte of data Per Flight, Says Virgin Atlantic. Accessed on 7 Dec 2016 [Online], Available : https://datafloq.com/read/self-driving-cars-create-2-petabytes-data-annually/172
13. Self-Driving Cars Will Create 2Petabytes of Data, What are the Big\Data Opportunities for the Car Industry? Accessed on 7 Dec 2016. [Online]. Available: http://www.computerworlduk.com/news/data/boeing-787s-create-half-terabyte-of-data-per-flight says-virgin-atlantic-3433595/
14. M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, The case for VM-base cloudlets in mobile computing. Pervasive Comput. **8**, 14–23 (2009)
15. ETSI, Mobile-edge Computing Introductory Technical White Paper, White Paper, Mobile-edge Computing Industry Initiative, (2014). https://portal.etsi.org/portals/0/tbpages/mec/docs/mobile-edge_computing_–_introductory_technical_white_paper_v1
16. Y.C. Hu, M. Patel, D. Sabella, N. Sprecher, V. Young, Mobile Edge Computing a Key Technology towards 5G, (2015). http://10.3.200.202/cache/8/03/etsi.org/6e14a9668574b8b935111768d9f6e501/etsi_wp11_mec_a_key_technology_towards_5g.pdf
17. ETSI GS MEC 001, Mobile Edge Computing (MEC) Terminology V1.1.1, (2016). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/001/01.01.01_60/gs_MEC001v010101p.pdf
18. ETSI GS MEC 002, Mobile Edge Computing (MEC) Technical Requirements V1.1.1, (2016). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/002/01.01.01_60/gs_MEC002v010101p.pdf
19. ETSI GS MEC 003, Mobile Edge Computing (MEC) Framework and Reference Architecture V1.1.1, (2016). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/01.01.01_60/gs_MEC003v010101p.pdf
20. ETSI GS MEC-IEG 004, Mobile Edge Computing (MEC) Service Scenarios V1.1.1, (2015). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/004/01.01.01_60/gs_MEC003v010101p.pdf
21. ETSI GS MEC-IEG 005, Mobile Edge Computing (MEC) Proof of Concept FrameworkV1.1.1,(2015). http://www.etsi.org/deliver/etsi_gs/MECIEG/001_099/005/01.01.01_60/gs_MEC-IEG005v010101p.pdf
22. ETSI GS MEC-IEG 006: Mobile Edge Computing Market Acceleration MEC Metrics Best Practice and Guidelines V1.1.1, (2017) http://www.etsi.org/deliver/etsi_gs/MECIEG/001_099/006/01.01.01_60/gs_MEC-IEG006v010101p.pdf
23. T. Taleb, S. Dutta, A. Ksentini, M. Iqbal, H. Flinck, Mobile edge computing potential in making cities smarter. IEEE Commun. Mag. **55**, 38–43 (2017)
24. Open Fog Consortium, [Online]. https://www.openfogconsortium.org/
25. OpenFog Consortium Architecture Working Group, OpenFog Architecture Overview White Paper, https://www.openfogconsortium.org/wp-content/uploads/OpenFog-Architecture-Overview-WP-2-2016.pdf

26. Q. Shen, L. Huang, G. Zhang, J. Gong, Policy Control and Traffic Aggregation for M2M Services in Mobile Networks. in *International Conference on Mechatronic Sciences, Electric Engineering, and Computer (MEC '13)*, pp. 3391–3395 (2013)
27. Cisco. Krikkit open source software, Feb 2014. http://eclipse.org/proposals/technology.krikkit/
28. H. Zou, Y. Yu, W. Tang, H.-W. M. Chen. FlexAnalytics: a flexible data analytics framework for big data applications with IO performance improvement. Big Data Res. J. **1**, 4–13 (2014). Special Issue on Scalable Computing for Big Data
29. R. Willett, A. Martin, R. Nowak. Backcasting: adaptive sampling for sensor networks. in *Proceedings* of the 3rd international symposium on Information processing in sensor networks. ACM, pp. 124–133. (2004)
30. A. Jain, E.Y. Chang, Adaptive sampling for sensor networks. in *Proceeedings of the 1st international workshop on Data management for sensor networks: in conjunction with VLDB 2004*. ACM, pp. 10–16, (2004)
31. S. Li, L. Da Xu, X. Wang, Compressed sensing signal and data acquisition in wireless sensor networks and internet of things. IEEE Trans Ind Inf **9**(4), 2177–2186 (2013)
32. X.-Y. Liu, Y. Zhu, L. Kong, Y.G. Cong Liu, A.V. Vasilakos, M.-Y. Wu, CDC: Compressive data collection for wireless sensor networks. IEEE Trans Parallel Distributed Syst **26**(8), 2188–2197 (2015)
33. F. Al-Turjman, Fog-based caching in software-defined information-centric networks. Comput Electr Eng J **69**(1), 54–67 (2018)
34. B.A. Bash, J.W. Byers, J. Considine. Approximately uniform random sampling in sensor networks. in *Proceeedings of the 1st international workshop on Data management for sensor networks: in conjunction with VLDB 2004*. ACM, pp. 32–39, (2004)
35. S. Lin, B. Arai, D. Gunopulos, G. Das, Region sampling: Continuous Adaptive Sampling on Sensor Networks, IEEE 24th International Conference on Data Engineering, (2008), https://ieeexplore.ieee.org/document/4497488/
36. N. Kimura, S. Latifi, A survey on data compression in wireless sensor networks. in *International Conference on Information Technology: Coding and Computing (ITCC'05)-Volume II*, Vol. 2. IEEE, 8–13, (2005)
37. Z. Huang, W. Lu, K. Yi, Y. Liu, Sampling based algorithms for quantile computation in sensor networks. in *Proceedings of the 2011 international conference on Management of data – SIGMOD '11*. ACM Press, New York, New York, USA, 745. (2011). https://doi.org/10.1145/1989323.1989401
38. K.-W. Fan, S. Liu, P. Sinha, Structure-free data aggregation in sensor networks. IEEE Trans. Mob. Comput. **6**(8), 929–942 (2007)
39. A. Papageorgiou, B. Cheng, E. Kovacs, Real-time data reduction at the network edge of Internet-of-Things systems. in *2015 11th International Conference on Network and Service Management (CNSM)*. IEEE, pp. 284–291. (2015). https://doi.org/10.1109/CNSM.2015.7367373
40. K.B. Pratt, E. Fink, Search for patterns in compressed time series. Int. J. Image Graph. **2**, 89–106 (2002)
41. J.W. Patty, E.M. Penn, Analyzing big data: social choice and measurement. Polit. Sci. Polit. **48**(01), 95–101 (2015)
42. M. Trovati, Reduced topologically real-world networks: a big-data approach. Int. J. Distrib. Syst. Technol. (IJDST) **6**(2), 13–27 (2015)
43. M. Trovati, N. Bessis, An influence assessment method based on co-occurrence for topologically reduced big data sets. in *Soft Computing*, pp. 1–10, (2015)
44. B. Jalali, M.H. Asghari, The anamorphic stretch transform: putting the squeeze on "big data". Opt. Photonics News **25**(2), 24–31 (2014)
45. B. Di Martino et al., Big data (lost) in the cloud. Int. J. Big Data Intell. **1**(1–2), 3–17 (2014)
46. P. Jiang et al., An intelligent information forwarder for healthcare big data systems with distributed wearable sensors. IEEE Syst. J. **99**, 1–9 (2014)
47. 2015 11th International Conference on Network and Service Management (CNSM) in [Real-Time Data Reduction at the Network Edge of Internet-of-Things Systems Apostolos

Papageorgiou, Bin Cheng, Erno Kovacs ¨ NEC Laboratories Europe Heidelberg, Germany apostolos.papageorgiou@neclab.eu, bin.cheng@neclab.eu, ernoe.kovacs@neclab.eu]

48. M.H. Rehman, P.P. Jayaraman, S.R. Malik, A.R. Khan, M.M. Gaber, RedEdge: A Novel Architecture for BigData Processing in Mobile Edge Computing Environments. J. Sens. Actuator Netw. **6**(3), 17 (2017). https://doi.org/10.3390/jsan6030017
49. D.Y. Kim, S. Kim, J.H. Park, A combined network control approach for the edge cloud and LPWAN-based IoT services. (2017). https://doi.org/10.1002/cpe.4406
50. C. Yang et al., A spatiotemporal compression based approach for efficient big data processing on Cloud. J. Comput. Syst. Sci. **80**(8), 1563–1583 (2014)
51. K. Ackermann, S.D. Angus, A resource efficient big data analysis method for the social sciences: the case of global IP activity. Procedia Comput. Sci. **29**, 2360–2369 (2014)
52. B.H. Brinkmann et al., Large-scale electrophysiology acquisition, compression, encryption, and storage of big data. J. Neurosci. Methods **180**(1), 185–192 (2009)
53. M. Weinstein et al., Analyzing big data with dynamic quantum clustering. arXiv preprint arXiv:1310.2700, (2013)
54. A. Cichocki, Era of big data processing: a new approach via tensor networks and tensor decompositions. arXiv preprint arXiv:1403.2048, (2014)
55. L. Zhang et al., Named data networking (NDN) project. Xerox Palo Alto Res. Center, Palo Alto, CA, USA, Tech. Rep. NDN-0001, (2010)
56. D. Raychaudhuri, K. Nagaraja, A. Venkataramani, MobilityFirst: a robust and trustworthy mobility-centric architecture for the future Internet. ACM SIGMOBILE Mobile Comput. Commun. Rev. **16**(3), 2–13 (2012)
57. F. DaCosta, *Rethinking the Internet of Things: A Scalable Approach to Connecting Everything* (ApressOpen, New York, NY, 2013)

Chapter 5
Applications of the Internet of Things with the Cloud Computing Technologies: A Review

U. Ram Jagannath, S. Saravanan, and S. Kanimozhi Suguna

5.1 Introduction to Internet of Things (IoT)

The term "Internet of Things" came into existence from 1999. Internet of Things has the power to change the world and is a system of interconnection between uniquely identifiable computing devices and mechanical and digital machines. This interconnection is made without involving human-to-human or human-to-machine interaction, and instead the data are transferred over the network. Internet of Things is a general concept where the devices could gather the data around the world and could share the data over the network where it could be utilized for various purposes. Wikipedia states, "The Internet of Things (IoT) is the interconnection of uniquely identifiable embedded computing devices within the existing Internet infrastructure"; i.e., it connects the embedded system to the internet. In recent years, there is an enormous development in the IoT field. Embedded systems are much attractive because of their low cost, low space, low power consumption, and autonomous factors. It also includes other factors such as I/O speed and cost, energy consumption, etc., based on the instructions.

Figure 5.1, obtained from Wikipedia, depicts the technological roadmap for Internet of Things. From this figure, it can be observed that by 2000, RFID tags were used to track the routing of the inventory to prevent the loss of logistics. This paved way for increase in the demand for logistics by providing support for the supply chain helpers. By the mid of 2000–2010, vertical-market applications such as surveillance, security, healthcare, transport, food safety, document mapping have

U. Ram Jagannath · S. Saravanan
School of Computing, SASTRA Deemed to be University,
Thanjavur, Tamil Nadu, India

S. Kanimozhi Suguna (✉)
Department of CSE, School of Computing, SASTRA Deemed to be University,
Thanjavur, Tamil Nadu, India

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in
Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_5

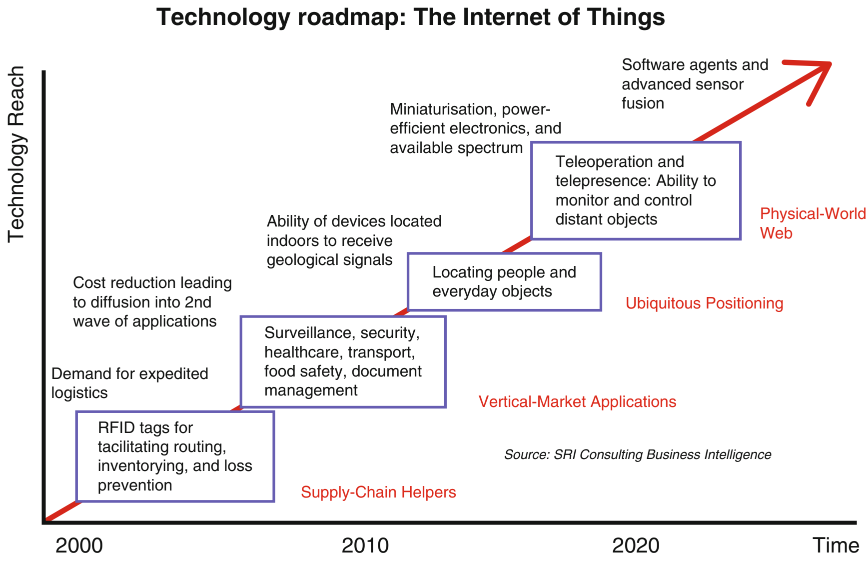


Fig. 5.1 Technology roadmap: The Internet of Things

seen a great development in technological aspects. This development in vertical-market applications has led to the progression of ubiquitous computing, in which the technology helps to locate people and everyday objects. Locating the objects that reside anywhere even indoors or undergrounds where satellite technologies may be unavailable or inadequate is called ubiquitous positioning. As a next stage, the progress is towards the physical-world web for miniaturization, power-efficient electronics and available spectrum to control distant objects.

Internet of Things could also be defined as the network of sensors where the data are exchanged under the connectivity using network protocols and systems. The exchange of data may be bidirectional or unidirectional where any one system could send or receive data or both.

The organizations or the field of applications which heavily rely on data will make use of cloud, fog, or edge computing. These computing architectures help in greatly utilizing the variety of computing techniques, resources that provide data storage, and these also include the Industrial Internet of Things (IIoT) [1]. Though these computing mechanisms sound similar, they form different layers in IIoT. Figure 5.2 is obtained from (www.winsystems.com); in this image, the authors have described the layers of IIoT. According to this site, cloud computing (or cloud layer) is treated as a major layer, and the extended layers of it are edge computing (or edge layer) and fog computing (or fog layer). These fog and edge layers are a collection of servers which includes distributed network. These layers help the organization by giving feasibility; hence, they need not keep the entire infrastructure on their own premises. This feasibility makes the data to be collected from any place, and it can also enhance the ability to access data allover. In order to increase the utilization of end-users and the IIoT devices, recent enterprises are moving towards edge or fog infrastructure.

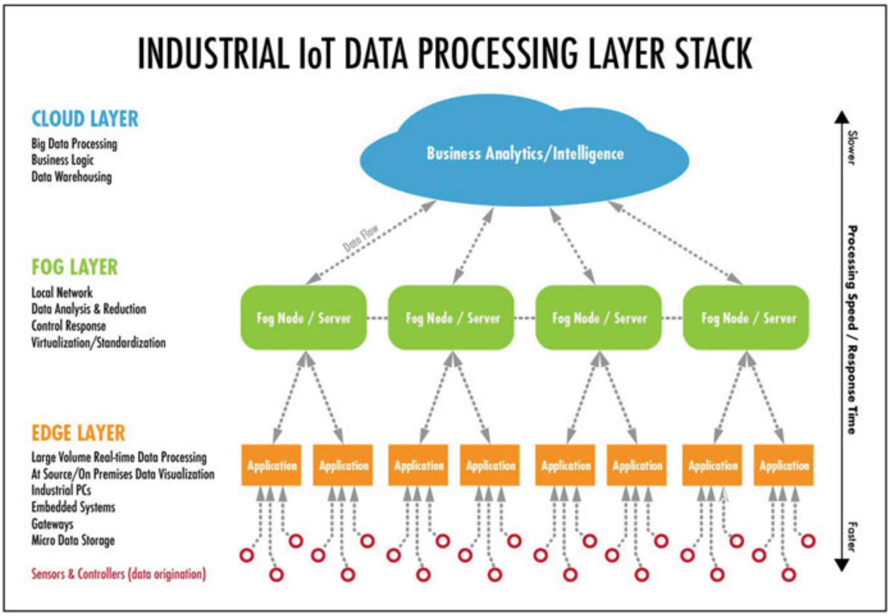


Fig. 5.2 IIoT data processing layer stack

Fog and edge computing look similar in terms of intelligence and also in the way the data are processed. But they differ in identifying the location of the intelligence and also in power computation. In edge computing environment, intelligence is placed in devices like embedded automation controllers, whereas fog computing environment places intelligence in the local area network (LAN) itself. For processing the power, fog computing uses edge devices and the gateways along with the LAN, but in the case of edge, the processing power is placed with the intelligence. Hence, the most challenging features in this computing are power consumption, latency, security, and also the data size which can be distributed across the network.

Older IoT technologies include a virtually infinite number of endpoints connected to the network. Because of this advantage, data collection and processing has become a biggest challenge. In order to solve this challenge, edge computing is introduced. The computation processes which are performed at the nodes have given a feel as if it is performed at the data source itself. Edge computing can also be stated as an extension of older technologies. The extended features available in the edge computing are peer-to-peer networking, distributed data, self-healing network technology, and remote cloud services. Improved hardware security and also lowering the power requirement are satisfying factors that are provided by the processors of edge computing systems. This edge computing system also provides a highly optimal solution as the hardware is embedded with the flash-storage arrays. The concept of interaction of the centralized system with embedded systems and industrial gateways is used in fog computing. Much of the data processing of edge computing is performed directly at the embedded computing platforms by making an interface with the sensors and controllers. Optimizing resource usage in a cloud-computing system is one of the advantages provided in edge computing besides more advantages like reducing network traffic and improved security. The following subsections explicate the IoT applications along with the concept of cloud computing.

5.2 IoT Applications

Table 5.1 gives a short description of the IoT applications which will be discussed in detail.

5.2.1 Infrastructure Management

IoT infrastructure consists of three major parts such as sensing infrastructure, computing infrastructure, and storage infrastructure. Sensing infrastructure includes communicating and sensing devices. Computing infrastructure is needed for computing the power consumption and to effectively manage the dynamic loads from various sensors. Depending upon the user need, this infrastructure has to be modeled. Traditional storage infrastructure has to be replaced with updated storage infrastructure in order to satisfy the users, as there is a drastic increase in the need of sensor. Infrastructure management is needed to monitor and control the devices in rural and urban area infrastructures. This infrastructure includes various applications such as bridges, railways, farms, etc. The infrastructure is controlled by monitoring the change in state of the application devices. If the device is compromised whenever there is a change in the state of an event, and if unnoticed will result in high risk. Infrastructure management helps in taking quick decisions and also in saving money with real-time analytics. It can also be used to schedule an event like repair and maintenance activities in industry, which will be performed by coordinating the task between the service providers and the users. With these facilities, IoT helps in improving quality of service, timely event handling process, coordinating the responses, and also reducing the cost. Many fields like manufacturing, agriculture, energy management, environmental monitoring, building and home automation, metropolitan scale deployments, etc, are benefitted with this infrastructure management.

Table 5.1 Short description of IoT applications

S.No.	Application	Description
1	Infrastructure management	Controls the operations of urban and mass infrastructures
2	Industrial IoT	Empowers industrial engineering with sensors and big data analytics to create brilliant machines
3	Smart retail	Makes automatic order of the supplies needed for the living
4	Smart supply chain	Notifies the receiver with details of the placed order
5	Transportation and logistics	Ensure less accident by providing more rules for managing vehicles
6	Connected cars	Bring in the concept of connecting the cars with the nearby places
7	Self-parking vehicles	Sense the sign symbols around the vehicle and flow to the parking lot according to the control
8	Smoke detectors	An environment that is contaminated by industrial emission is controlled by using it
9	Smart cities	An eco-friendly environment with less pollution and pollutants
10	Constructions	Automated machines to construct buildings by itself for our well-being
11	Smart homes	Assisting household works with the help of IoT
12	Smart panels	Contamination of water and other things are saturated using such devices
13	Energy management	Consumption of electricity is majorly sensed and stocks the extra
14	Smart wearables	Get the important appointments done by reminder which helps in smart living and also reduces industrial accidents
15	Environment monitoring	Senses to assist with human life and wildlife with natural disaster warnings
16	Biological sensors	Used to study DNA and other molecules
17	Healthcare systems	Read the body and suggest the health measures to be added
18	Precision agriculture	High demand for quality food is met with the growing field of IoT
19	Poultry and farming	Livestock monitoring is about animal husbandry and also is cost saving
20	Smart surveillance	Stealing personal information has an issue with human and hence assures security from thieving

The main idea for growing applications is cited by Li Da Xu, Wu He, and Shancang Li, “Internet of Things in Industries: A Survey”

5.2.2 Industrial IoT (IIoT)

IIoT describes the usage of Internet of Things across several industrial activities like manufacturing, logistics, metals, and mining and in other sectors. Internet of Things is a transformative manufacturing technology which improves safety, quality, and productivity in industries [3]. The main aim of the Industrial Internet of Things is to have a high output at the rate of same input and to achieve same output in the cost of low input. All goods and products in the industry are labeled with near field communication (NFC) tags and radio-frequency identification (RFID) tags which helps to monitor the goods available in the industry and helps to order goods if it is insufficient. It also helps to reduce time in searching the specific products in the industry and also reduces labor cost.

The primary advantages of the Industrial Internet of Things are as follows:

- Predictive maintenance
- Real-time monitoring
- Scalability and time saving
- Resource optimization

The study, The Impact of Connectedness on Competitiveness, is based on a global survey of 350 executives from large enterprises. It found that more than half (52%) of executives at large enterprises expect IIoT to have a "significant" or "major" impact on their industry within 3 years, with another third (32%) forecasting a "moderate" impact in that timeframe. Security is the major factor for all industries and everyone is in need to know that their data is well secured. Security vulnerabilities are highly prevented by using sensors, smart connected devices, and by other methods.

5.2.3 *Smart Retail*

IoT brings a number of applications for retailers for improving store operations, minimizing theft occurrences; increasing customer purchase by selling more number of the products to them, and also producing a new inventory product [7]. It will make the customer know everything in the store and aid them closer to the buying decision. RFID tags are attached to every product in the store, and its availability is checked periodically. Customers gain the information about the goods through Bluetooth beacons or other digital receivers. The digital signature will navigate the customer to the product. The following actions can be performed for improvising the quality of purchase and make the customer think and take an easy decision in buying a product. From the analytics, the retailer will know about customer's perspectives which will later help in improvising their product according to the customer's interest.

- Interactive digital signature screens
- Mobile shopping applications
- Sensor-based items tracking
- Tailored digital marketing
- Optimized inventory management

- Mobile payment solutions
- IoT-enabled beacons
- Customer preferences analytics

For security purposes in the retail shops, multiple video cameras are positioned within the shop premises at present. Using these cameras, images are captured and are streamed to the cloud for storage and also for further analysis, if required. These are implemented for security reasons or can also be used to monitor the customer behavior within the store. As an extension, it helps the proprietor of the store to keep an eye on the happenings of the store including the cash flow. For implementing a secured store, the store is in need of at least one gateway to connect various equipment with the sensors and also to collect the image data and store it in the cloud platform. Many companies are working on this issue to give better security to the retail store.

5.2.4 Smart Supply Chain

IoT supports supply chain management by performing operational efficiencies and revenue opportunities. The supply chain transparency captures the extent to which information about the companies, suppliers, and sourcing locations are readily available to end-users and also to other companies in the supply chain. Security is the most important aspect that has to be considered. All the information must be prevented from falling into the wrong hands or being hacked. Sensors should only send specific information, which must be held in a secure private cloud environment.

Product Tracking RFID and GPS sensors track product from the manufacturing place to the storage location. They will monitor the product, and hence the manufacturer can gain granular data like the temperature at which an item is stored and so on. These data will help in improvising the quality control, on-time deliveries, and product forecasting.

Vendor Relations Data obtained from product tracking are also helpful to companies for updating their own production schedules and also recognize the vendor relationships that may be costing them money. These will test how vendors are handling supplies and sending goods to the manufacturer. High-quality goods define a better relationship with customers, and the better customer is retained.

Forecasting and Inventory Another major parameter which is difficult to manage in the supply chain management is inventory. Smart devices are closely monitored, and they deal with more variable than earlier applications. Equipment monitoring provides information about predictive maintenance and alerting engineers for potential breakdowns and costly disruption. Consider Amazon Go as an example that uses IoT, computer vision, machine learning, and AI to create a shopping experience where the user can just walk in, pick up whatever the user needs, and can walk out with a purchase.

In [15], the authors have proposed a cloud-centric IoT architecture for supply chain management (SCM). In traditional SCM, the movement of physical objects was performed directly, whereas, in this paper, the authors have controlled and coordinated the supply chain process based on object sensing. For implementing this framework, they have used proposed layers in their framework, namely, object management, cloud computing platform, information as a service, and smartphone user. Thirty-one percent of manufacturers are implementing Internet of Things enhancements to their internal operations, with another 56% exploring doing so in order to cut operational costs, achieve supply chain efficiencies, and improve predictive maintenance capabilities. Operations performed with these layers include submission of request by the user to the cloud; the cloud computing platform will make response to the user by identifying the location; this cloud computing platform helps the user to update the data in the database and hence the user can publish the sensed information; with the received information, the cloud will perform operation and will return the user about the task which is mapped according to the physical location of the user; and finally all the transactions and the conversations are stored to the cloud for future references.

5.2.5 Transportation and Logistics

IoT improves the transportation industry by improving economic development, public safety, and the environment. IoT monitors the goods and analyses the data with the help of sensors. Location of the vehicle can be predicted by GPS, roadside sensors, etc., and finally, the products will be delivered on time, at the right place. The real-time traffic monitoring is explained in [10]. Some special delivering vehicles are available to deliver sensitive products, to be stored in a cool place. Environmental sensors can monitor temperature, humidity, light, and vibration in real time. These sensors have the capability to detect the occurrence of any problem, and they are alarmed so that corrective measures can be taken to prevent spoiling of goods. The data that are collected from different sensors are massive in size to be stored in the cloud. Safety and security is the main reason for the development of the transportation industry with IoT. Thus, there is no loss of goods, and hence they are transported to the destined place without any damage. By maximizing the profitability and viability, the efficiency of the supply chain is improved in transportation and logistics field. With the improved development in the fields of mobile technology and IoT, the enterprises are connecting the transportation- and logistics-related devices with the centralized cloud network to gain information about the real-time operations in transportation and logistics by capturing and sharing the information about the critical data. The information about the critical data is shared in to make sure that the right product is delivered to the right place at the right time. This paves way for the development in various fields such as asset management, cloud, mobile, and big data.

5.2.6 Connected Cars

Internet of Things paves a great and fresh way for the car manufacturers to upgrade their cars in a highly competitive market. As stated in [2] by Kevin Ashton, this also enhances the idea, where a person can drive around the world and can stay connected all the time. Before the invention of IoT in automobiles, one has to search for the route manually, while the IoT has made it more convenient. For example, a person could set the destination place using Ground Positioning Systems (GPS), and the driver will be directed to the destination place and this also needs some improvements.

Auto-pilot mode is the most modern technology introduced in recent times. High-resolution cameras can be used to detect soft and hard objects in the lane and also in the vehicle surroundings. Forward-facing radar with enhanced processing collects additional data about the world on a redundant wavelength that helps the user to see through heavy rain, fog, dust, and even the car behind and ahead. This could also aid in detecting the traffic signals and signs that are already inbuilt in the system and could drive the car as safe as possible. It can also adapt to the changes in the environmental roads around the world, and it could also offer lane change automatically. IHS Markit predicts that more than 70 million connected cars will be on the road by 2023.

Designing a connected vehicle platform on cloud IoT core will be a better solution for managing the connected cars by making a use of cloud IoT core on Google Cloud Platform (GCP). The use cases that can be defined by making a combination with data storage and platforms are usage-based insurance, predictive maintenance, freight tracking, and customized in-vehicle experience. The main challenges that a development team will face in this domain are device management, data ingestion, data analytics, applications, and predictive models. Besides these challenges, the developer must also focus on security and set the boundary level for the applications they design that are implemented in the cloud-based system.

5.2.7 Self-Parking Vehicles

Self-driving cars are learned from the data cloud, the collections of data are brakes, tire and transmission performance, fuel consumption, and efficiency of the vehicle. The main reason for self-driving is high safety, and it will abide all traffic rules and regulation as shown in [5]. Averagely, parking coverage takes 31% of land use in big cities, like San Francisco, and even more, 81% in Los Angeles and 76% in Melbourne, while at the lower end we find New York (18%), London (16%), and Tokyo (7%). Hence, the removal of unnecessary parking areas is the first task for municipalities to create a better urban planning. The data are provided with high security, and the hackers cannot access the data easily without the user's authorization. There exists confusion when the user tries parking his or her vehicle, and it leads to great pressure and stress to the user which also results in waste of time. As self-driving vehicle is automatically parked in the parking area without any confusion it reduces pressure as well as the time for user. IoT predicts the current scenario and performs the action according to the problem.

In [16], the authors have designed a smart parking system. In this system, the authors have specified that the system can be implemented for any type of parking like parking available at covered parks, open parks, and also street side parking. For this system, the authors have designed an architecture which includes cloud service provider. The main function of this cloud service provider is to collect information about the parking area and store the information in the cloud. The major components of this smart parking system are centralized server, Raspberry Pi, image capturing device, a navigation system, displaying device, and the user device. The centralized server acts as a database that will hold the entire information about the parking system, some of the information includes a number of parking slots, availability of vehicles, and so on. Raspberry Pi is a microcontroller used to implement the parking system with a camera attached to it. This Pi-camera is used to capture pictures of the parking area for validating the area to check the availability of slots in that area. The main role of the navigation system is to navigate the users to the available parking lot which is nearer to his or her current location. Displaying device is a monitor or tab available at the admin's location to monitor the device and to modify the parking lot if needed. The user device is used to connect the user with the parking system using their smartphones or browsers.

5.2.8 Smoke Detectors

Safety of life is more important when compared to any other application. Records show that many lives were exploited by fire accidents, and also it takes a longer duration for the fire service to reach the destination. To avoid these types of accidents, many companies are using IoT devices for safety. It is safe in many ways as it is more evident that if fire accident occurs, it automatically connects itself to the hot-connect or cloud environment and it directly sends a voice or text message to the concerned fire agency. The working principle is not like the normal smoke detectors which could only give beeping alarm sounds but it also helps to save people. This has battery power that is estimated to last up to 10 years and also it is not a big bottleneck to change the battery. It is more rapidly replacing the normal smoke detectors as this proves to be more useful and also available at an affordable cost. One thousand smoke detector surveys were mailed, covering about 70 percent of the town's households. While the results staggered in over a period of 3 months, 315 completed surveys (31.5 percent) were ultimately returned. The high return rate can be interpreted as a strong interest on the part of the public concerning fire safety.

In the industrial sector, the smoke detectors play a different role, where it is used as the contamination detector. The detector is placed at the chimney and monitors the smoke that is being passed out through it. If the contaminants are greater than the recommended level, it will send a warning message to the main supervisor and also to the concerned person so that it will never be ignored. It ensures green living to all the people and greatly reduces pollution.

5.2.9 Smart Cities

As described by Gauer, A., Smart City Architecture and its applications based on IoT in [4], smart cities improve the quality of lifestyle by improving infrastructures like delivering clean water, dependable power, and efficient public lighting. The main motive of a smart city is to provide more efficient water supply, an innovative solution to traffic congestion, and more reliable public transportation. Water can be highly saved by providing the following facilities.

- Leakage detection in pipes
- Wastewater recycling
- Managing the storm-water

The benefits of the smart city include increased safety, reduced traffic, lower levels of pollution, more efficient use of energy, and improve the overall quality of life for future city dwellers. The main aim of the smart cities is extended to keep the citizens safe, and it should provide education and technological job opportunities to the citizens with open innovation. Gujarat International Finance Tec-City stands as a great example for smart cities. It has better city planning and development, and internet facility is provided all over the city with a faster supply of products at a lower operating expense. The major issue stands in the contamination of air, and to resolve this issue many green buildings are set up.

The companies which focus on security management can be connected with smoke detectors so that they can protect the buildings from fire. The device will send the message to the cloud-based dashboard only if the fire is detected or if the system is in need of maintenance. This will greatly reduce the man-hours in performing the maintenance check frequently, and even the process may not be needed. These connected smoke detectors are more advantageous when compared to the traditional system. These smoke detectors can be connected to devices such as desktops, laptops, tablets or smart-phones to get alert messages which proves helpful for fire security personnel.

5.2.10 Construction

The main problem faced by a worker in a construction industry is misplacing the equipment at the construction sites which leads to time consumption in searching the equipment or may cause any damage to the equipment. IoT keeps on monitoring the equipment by sensors by GPS tracking attached to the equipment. RIFD tags will be attached to the devices so that tracking can be made easy. The cost of the equipment is comparably higher than earlier devices. This will help the consumer to buy the equipment only once in a lifetime, instead of purchasing frequently whenever it is missed. Repair and service of the equipment can be monitored, and the problem could be resolved with a warning alert from the equipment. Energy consumption by the equipment can be minimized by the sensors which will adjust the actions to be performed by the tools automatically. The sensors will automatically switch on/off to save energy. The main advantage is the quality work is increased and more convenient for users. The construction project work can be done on time, and the quality of the construction can be improved. The future Internet of Things will provide intelligent building management systems which can be considered as a part of a much larger information system used by facilities managers in buildings to manage energy use and energy procurement and to maintain efficient building management systems.

5.2.11 Smart Homes

Google is highly potential and develops automation in everything. As everything in the home is well organized and fully computerized, they are all controlled simply with a mobile application which leads to controlling and monitoring the devices that are installed in the home [17]. The devices that are using electricity can be added to the network. Smart homes also provide high security which prevents hackers from penetration.

Several companies have developed smart appliances which include smart fridges, televisions, and air conditioners. For example, Nest is an automated thermostat which senses the temperature and adjusts the air conditioner to the favorable means, i.e., if the external temperature is high, it lowers the internal temperature and vice versa which makes the people comfortable. In [12], the authors describe the privacy and security challenges in smart homes.

In [18], the authors have designed a smart home in such a way that the device will record the daily activities inside the home and store the information in the cloud storage. The device will get familiar with this information, and hence at a later stage, the device will do the similar operations without the requirement of human intervention.

5.2.12 Smart Panels

A smart panel can perform multiple functions like tracking the electrical usage and wastage in a detailed manner, and it also displays the result on the computer screen. They are simple, flexible, and have an open architecture which proves that they are simple in construction and implementation. Using these applications, the user can remotely access the electrical units which are installed far away which help in reducing the electrical energy.

Another type is smart solar panels, and these have multiple abilities which could perform different actions depending on the conditions required. For example, this panel can be used to convert solar energy into electrical energy or to the required form of energy in which it acts as a transducer in one form and it could convert air into drinking water from the sunlight and air in another form. All the installed panels are monitored by the organization. The purity and taste of the water are obtained by adding the required minerals to it. It is connected and so it provides optimized performance. It could provide an average of 8–20 bottles of water per day. The main advantage of using these panels is they need not be hooked up to the home's electrical systems, and they use full renewable resources to gain energy.

5.2.13 Energy Management

Energy is a very important aspect for any household, industries, agriculture, and so on. Managing energy source and conserving it for better performance is a challenging task. IoT paradigm promises to increase the visibility and awareness of energy consumption, thanks to smart sensors and smart meters at the machine and production line level. Since every appliance depends on energy supply, energy plays a vital role in day-to-day life. An interface like Arduino Microcontrollers controls the usage of an appliance like the intensity of light and the speed of the fan. Humidity, temperature, and light intensity are taken into consideration in developing a better IoT system for energy management. IoT devices are controlled and connected for small cities and communities, transportation systems, buildings, lighting systems, factories, and more devices. Building Energy Management Systems (BEMs) attracted $1.4 B in VC Funding from 2000 to 2014 (26% of all investment in building energy technology). In 2020, about 77% of the $2.14 billion US market will comprise BEMS applications, and 40% will come from buildings below 50,000 square feet. The US market for sensors and controls for BEMs will rise at a 17% compound annual growth rate to $2.14 billion in 2020. Possible future directions for energy management are as follows:

- Energy-efficient mechanisms for software-defined IoT solutions, which can provide scalable and context-aware data and services.
- Fog computing can lead to energy saving for most of the IoT applications; therefore, it is important to study energy consumption of fog devices for IoT applications.

5.2.14 Smart Wearables

Wearable devices are considered as the hot topic these days. A reduced job is much preferred in every field which resulted in this technology. Some of the important functions to be considered for a smart wearable in the industrial location is it could sense cut-off radiation level and will alert the user when the environment becomes worse.

Health- and fitness-related devices are also available these days which could monitor the heartbeat rate, blood pressure, footsteps count, sleep monitor, etc. Some of the useful features are it could find the body temperature by which the user could get an early indication of cold or flu. Most of the users have a question why to use a wearable gadget while a single smartphone could do all these work. It is not comfortable to take our smartphone for every transaction, but the smartphone will act as the central unit of the wearable by connecting it through Bluetooth or the network which grants access for the wearable to surf the data available in the smartphone.

Further, wearable reduces manual work to a great extent where the devices in the home are being connected and the user can control any of the devices which are connected in the same network without requiring any movement from place to place.

5.2.15 Environmental Monitoring

The main application of IoT in the environmental monitoring is weather forecasting and monitoring, protecting endangered species, water safety, and many more applications. Current weather forecasting and monitoring applications include the need for the labor work, but implementing IoT in this field greatly reduces this effort. Sampling is done at regular intervals, and it helps to prevent contamination of air in many stages. IoT allows deep monitoring of weather in a specific range of area using IoT installed with radar. This promises more fine-grained data with great accuracy and range. This helps in early detection and early warning which reduces the loss of life and property [11]. Low-power, wide-area network (LPWAN) technology is perfectly suited for environmental monitoring, as it can connect devices that need to stay in the field for an extended period of time and send small amounts of data over a long range. There are a number of reasons you may want to select LPWAN technology for your IoT environmental monitoring:

- Long battery life
- Low cost
- Long range
- Satellite backhaul ability

5.2.16 Biological Sensors

Biosensors are gaining more interest nowadays, and they promise to be one of the constant growing fields in IoT [9]. Biosensors are mainly used in the field of healthcare, applications related to sports, and mostly in the military. The main advantage of this application is low cost and also provides real-time information. The term "wearable biosensors" (WBS) is evolved by combining both the smart wearable and the biosensors in a single unit. The required sensors are installed in the various parts of the body so they could be monitored all the time, without any whatsoever disturbances. Their size will be very small which makes them easy to carry wherever the user moves. Based on the usage of such sensors, these devices are popularly named as immunosensors, glucometers, biocomputers, etc., some of the examples include (1) Google lens (2) ring sensor and (3) smart shirt.

5.2.17 Healthcare Systems

According to the World Health Organization (WHO), 31 percent of global deaths are due to cardiovascular diseases (estimated 17.7 million people in 2015). The number of people with diabetes has been rapidly increasing from 108 million in 1980 to 422 million in 2014. The rising cost of healthcare services has increased the need for providing effective and efficient healthcare to the patients in most of the developing countries. Internet of Things stands as the key component for the digital transformation of healthcare systems. Internet of Healthcare Things (IoHT) and Internet of Medical Things are considered to be the two most important parts of healthcare systems which include smart pills, smart personal healthcare, and real-time health systems. This is mainly designed for those who do not find time to spend on their health life, whereas here they could easily check their health issues with easy IoT healthcare systems. This is not only suitable for an individual but easy for everyone to register with their health issues and get appropriate remedy, and also this initiative will reduce spending money for a periodical check-up.

IoT-based healthcare for the elderly is the cheapest and more efficient way for applications related to healthcare. ECG, heartbeat rate, and body moisture can easily be found using this application. Appointment of the doctors can be made and it will also help in periodical indication for the pills at right time. Healthcare IoT systems ensure some basic ideas that can now be used to monitor the patient remotely, but in future, it will surely be an important member in the healthcare industry to maintain the healthcare and its related activities. The health records of the users are collected and are stored in electronic healthcare records (EHR) which makes the past data as well as the present data to be readily available at any time. While considering the security of the personal health records, IoT considers this as sensitive data and has some stricter rules that should be followed. Genetic materials and other personal details are stored in a private environment.

In [13], the authors have proposed a device named Cellcon. This device is developed based on the image analysis concept which is portable and also affordable to do the process. This is a hardware device with a ball lens and a mobile application. The authors designed this application to do the cell count and to identify the presence of the malarial parasite. The blood sample is collected in a slide and a photo is taken by a ball lens that is attached to the phone clip, and hence a magnified image of the blood sample is obtained. This IoT device uses cloud storage for storing the image data, retrieve the data, and perform the analysis through and from the internet. The image is compared with the images in the database and the results are sent to the patient. In this application, the first half of the application, i.e., collecting the blood sample and taking a photo, is performed at the patient's location. The patient may reside at any remote location, and the image data will be passed via the internet. The doctor will analyze the image data using the application which is installed in his mobile, and this belongs to the second phase of the application.

5.2.18 Precision Agriculture

Precision agriculture makes rapid changes in the agricultural sector, and numerous organizations are leveraging this technique around the world. IoT makes agriculture field highly efficient. CropMetrics is a precision agricultural organization concentrated on ultra-modern agronomic solutions. Wireless sensor network (WSN) and wireless moisture sensor network (WMSN) are used in IoT. WSN technology contains many sensors, which are used to monitor and control to do proper irrigation. This optimization will maximize the profitability of irrigated crop fields, improves yields, and increases water usage efficiently. Agricultural drones are applied to farming in order to help increase crop production and monitor crop growth. Through the use of advanced sensors and digital imaging capabilities, farmers are able to use these drones to help them gather a richer picture of their fields. An advantage of drones is crop health, convenient to use, and increase in yield production. IoT reduces the human intervention in farming. Today's sophisticated commercial farms have exploited advanced technologies; however, IoT introduces more access to deeper automation and analysis.

In [14], the authors have projected the problems faced by the farmers in the real world. The farmers will be holding the agricultural land which is measured in hectares, and it is highly difficult for the farmers to identify the infected crop that is spread out in a vast area. Besides these criteria, the farmers also face certain other issues like rainfall prediction, climate, availability of water, and so on. Also, the farmers are not ready to cultivate different or hybrid crops in their land, as they lack knowledge and less experience in this scenario. In order to support the farmers in these conditions, a sensor device can be placed in the soil that is used for monitoring the climatic changes, temperature, and even soil moisture. The data observed by the sensors are stored with the help of cloud storage; this stored information will be used in future for analyzing and predicting the soil status and also about the feasible crop which can be cultivated with the existing soil and weather conditions. This concept of implementing cloud will aid in the expansion of data storage, and this will greatly reduce the production cost.

5.2.19 Poultry and Farming

IOT is an innovative technology that provides modern and automated poultry farming. Basic environment parameters for poultry farming are temperature, humidity, ammonia gas, and water level. For instance, the temperature is controlled by cooling fan, humidity is monitored by the exhaust fan, ammonia gas is examined by ventilation window open, and the water level is maintained by DC motor. These actions are performed by the sensor, and monitoring is done the whole day. These are done mainly for the growth and health of the chicken. Farming challenges caused by population growth and climate change have made it one of the first industries to utilize the IoT in this sector.

5.2.20 Smart Surveillance

The major issue in the secured environment is to safeguard our buildings, assets, and so on [8]. The basic alarm facilities of earlier days had high visibility and later on introduced a monitoring camera which has some less scope of treatment. Internet of Things helps to create smart homes and smart cities by adding a feature of real-time monitoring services to the environment which concerns more about the security by the use of smart surveillance and security [6].

To make it simpler, let us consider that the user uses network cameras to monitor the area under surveillance and high-resolution cameras for the places of large crowd movement. IoT will offer a greater solution to such disparate things by considering the above scenarios by combining the video surveillance cameras and smoke detectors into a "small and single glass pane." As all the devices will be monitored over a network, they could share data among them and could take decisions by themselves by which the security will be enhanced.

The images produced by the video surveillance cameras should have high clarity, and this would be true by adopting Internet of Things. Nowadays, megapixel technology has grown to its peak by incorporating them with low-lighting cameras which are now blooming successor of the present market. These are all concerned with the wide dynamic range (WDR) of cameras. IoT offers a resolution of at least three times greater, and even 4 k videos could be generated using IoT.

5.3 Conclusion

There are many benefits that arise with the increased use of Internet of Things. IoT is used in our everyday life as it greatly reduces the cost and power. Hence, this chapter has given an overview of applications using IoT. IoT has the potential to drive integrated solutions that can make a difference. The entire world is transforming slowly with the use of IoT in daily life. Also, the people are getting ample opportunities to enhance their commitments and perform various activities using smart things. This chapter has given a short description of various applications that are implemented with IoT and also certain challenges faced in those applications. Research work can be extended in the challenges faced and giving increased security to the applications to provide the better, smart and secured environment.

References

1. L. Da Xu, W. He, S. Li, Internet of things in industries: a survey. IEEE Trans. Ind. Inf. **10**(4) (2014)
2. K. Ashton, That 'internet of things' thing. RFID J. **22** (2009)
3. D.Y. Joo, J.K. Kim, Creative & active convergence model of IoT. Korea Inst. Ind. Econ. Trade Korea **1** (2014)
4. A. Gauer, Smart city architecture and its applications based on IoT. Procedia Comput. Sci. **52**, 1089–1094 (2015)
5. A. Bagula, L. Castelli, M. Zennaro, On the design of smart parking networks in the smart cities: An optimal sensor placement model. Sensors **15**(7), 15443–15467 (2015)
6. K. Sohraby, D. Minoli, T. Znati, *Wireless Sensor Networks: Technology, Protocols, and Applications* (Wiley, Hoboken, 2007)
7. H. Sundmaeker, P. Guillemin, P. Friess, S. Woelfel, *Vision and Challenges for Realizing the Internet of Things* (Publications Office of the European Union, Luxembourg, 2010)
8. B. Khoo, RFID as an enabler of the internet of things: issues of security and privacy, in *Internet of Things (iThings/CPSCom)*, 2011
9. Y. Zhang, Technology framework of the internet of things and its application, in *Electrical and Control Engineering (ICECE)*, 2011
10. Y. Cao, W. Li, J. Zhang, Real-time traffic information collecting and monitoring system based on the internet of things, in *Pervasive Computing and Applications (ICPCA)*, 2011
11. A.R. Jaladi, K. Khithani, P. Pawar, K. Malvi, G. Sahoo, Environmental monitoring using wireless sensor networks (WSN) based on IOT. Int. Res. J. Eng. Technol **4** (2017)
12. J. Bugeja, A. Jacobsson, P. Davidsson, On privacy and security challenges in smart connected homes, in *European Intelligence and Security Informatics Conference*, 2016
13. Z.H. Chan, E.L.M. Su, Y.M. Lim, E.E. Langeh Randin, C.F. Yeong, M.A. Abdul Razak, I. Ariffin, Cellco: portable device for automated blood cell count and abnormal cell detection. J. Telecommun. Electron Comput Eng **9**(3–9) (2017)
14. B.B. Prahlada Rao, P. Saluja, N. Sharma, A. Mittal, S.V. Sharma, Cloud computing for internet of things & sensing based applications, in *6th International Conference on Sensing Technology (ICST – 2012)*
15. P. Satpute, A. Mohanpukar, Cloud centric IoT framework for supply chain management. Int. J. Comput. Appl. **118**(15) (2015)
16. S.R. Basavaraju, Automatic smart parking system using internet of things (IoT). Int. J. Sci. Res. Publ. **5**(12) (2015)

17. J. Rajaleksmi, G.G. SivaSankari, IoT framework for smart home using cloud computing via open source mobile platform. Int. J. Comput. Eng. Appl. (2016), ICCSTAR-2016, Special Issue, May 2016
18. M. Kasmi, F. Bahloul, H. Tkitek, Smart home based on internet of things and cloud computing, in *7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT'16)*, 2016

Chapter 6
Software-Defined Internet of Things to Analyze Big Data in Smart Cities

Sadia Din, Awais Ahmad, Anand Paul, and Gwanggil Jeon

6.1 Introduction

With the use of the ICT, which is one of the fascinating technologies, not only people are connected with each other but also now things are connected with everything. Accessibility to all kind of services is made so easy. Health, transportation, emergency response, shopping, utilities, economy, weather, etc., are the main services which are available through the ICT. Some researchers, including the authors of this chapter, referred these services as smart services. Smart cities concept also originated because of advancements in ICT research and inventions. The generation of data is very high these days as IBM reports that approximately 90% of the data is generated in the last 2 years [1]. Many research organizations have now started research on big data and related topics because of this fact. Big data consists of large sets of data which can be mainly identified through "5Vs" (initially, it was referred to as "3Vs"; two have been added recently): volume (size of dataset), variety (range of data type and source), velocity (speed of data in and out), value (how useful the data is), and veracity (quality of data) [2]. Because of its complexity in processing, big data becomes more challenging. The conventional network obtained great success by adopting a hierarchical paradigm. Anyway, the closed systems of network devices in smart cities, we have to configure many devices with high

S. Din · A. Paul (✉)
School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea

A. Ahmad
Department of Computer Science, Bahria University, Islamabad, Pakistan

G. Jeon
Department of Embedded Systems Engineering, Incheon National University, Incheon, South Korea

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_6

complexity when business requirements change. The researchers also cannot deploy new protocols in the real environment especially when we are focusing on a plethora of embedded devices.

It has been expected that the internet growth will increase to very high end, that is, 1.6 × 1021 B [3]. In order to cope with this challenge, a greater bandwidth will be required by users so they can access any services with good speed. The stability of network will also be required. Almost half decade ago, the abovementioned issues led Nick McKeown to present Open Flow [4] and based on Open Flow, the concept of software-defined networking (SDN) was presented in [5]. Where SDN is based on the idea of detaching the control plane from the forwarding plane, to break vertical integration, and to introduce the dynamic network and with the ability to program it. The SDN allows logical centralization of feedback control, and decisions are made by the "network brain" with a global network view, which eases network optimization. The data plane elements become highly efficient and programmable in the SDN system with packet forwarding devices, and the control plane elements are represented by a single entity, the controller. In conventional networks, it is much easier to develop and deploy applications in SDN. Also, with the global view in SDN, it is straightforward to enforce the consistency of network policies. SDN represents a major paradigm shift in the evolution of networks, introducing a new pace of innovations in networking infrastructure.

6.2 A Brief Overview of the SDN and Big Data Roles in the Development of Smart Cities

Three main advantages of the SDN architecture are as follows: (1) The SDN's open architecture guarantees the centralized control and automatic management of networks. SDN allows network engineers to design, deploy, operate, and maintain networks using a centralized SDN controller rather than configure a lot of heterogeneous devices. (2) The SDN can provide various open APIs to flexible program networks. The network operating system and network applications can be deployed on servers that adopt X86 architecture and can control data forwarding by Open Flow. (3) The SDN architecture decouples the data plane and the control plane by using Open Flow and virtualized networks. A network becomes a logical resource that can be configured through software. The core idea of SDN has been used in the field of routers to build an open, flexible, and modularized reconfigurable router [6]. While some excellent work has been done on big data and SDN, these two important areas have conventionally been addressed separately in most of the previous works. However, on the one hand, SDN, is an important networking paradigm, will have a significant impact on big data applications. Several good features, for example, separation of the control and data planes, logically centralized control, global view of the network, and ability to program the network, are associated with it. For example, big data is usually processed in cloud data centers. Compared to traditional data centers, SDN-based data centers can have better performance by dynamically allocating resources in data centers to different big data applications to meet the service-level agreements (SLAs) of these big data applications [7–19].

On the other hand, big data is an important network application which has a profound impact on the design and operation of SDN. Specifically, with the global view of the network, the logically centralized controller in SDN can obtain big data from all the different layers (i.e., from physical to application layers) with arbitrary granularity.

Experience shows in cross-layer design that sharing information among different layers can improve network performance, the network becomes so complex that conventional methods are unsuitable to design and optimize such networks. Luckily, big data analytics, which leverages analytical methods to obtain insights from data to guide decisions, can help the design and operation of SDN, for example, with big traffic data analytics, it is easier for the controller to perform traffic engineering to improve the performance of SDN.

This research proposes the integration of programmable devices with IoT (SDN-IoT) to provide a variety of big data applications for smart cities followed by the potentials of SDN in building smart cities. The proposed paradigm starts with data gathering process from various smart city–enabling technologies such as smart homes, smart grids, intelligent transportation system (ITS), weather forecast intelligent systems, and so on. After the data aggregation as a second level, we present a detailed data processing and management (DPM) level where big data analysis has been used to filter the useful data. We also have evaluated the performance of our architecture using Hadoop ecosystem. On top of DPM, we have also identified the applications level along with varying data flow that is supported by our evaluated big data architecture. In this, we have considered the future internet architecture recently named as named data networking (NDN) also known as content-centric networking (CCN) at the application level. This chapter, finally, provides open issues that provide a roadmap to follow by the active researchers in the said domain.

6.3 Proposed Scheme System Architecture for SDIoT

This section presents details of the proposed novel layered architecture which support high-performance computing for SDIoT and also help to design the system completely. The proposed system architecture consists of four multilevel with the objective to enhance the current capacity of storing and the speed of processing. At each layer, various tasks are supported by the reading and writing capabilities.

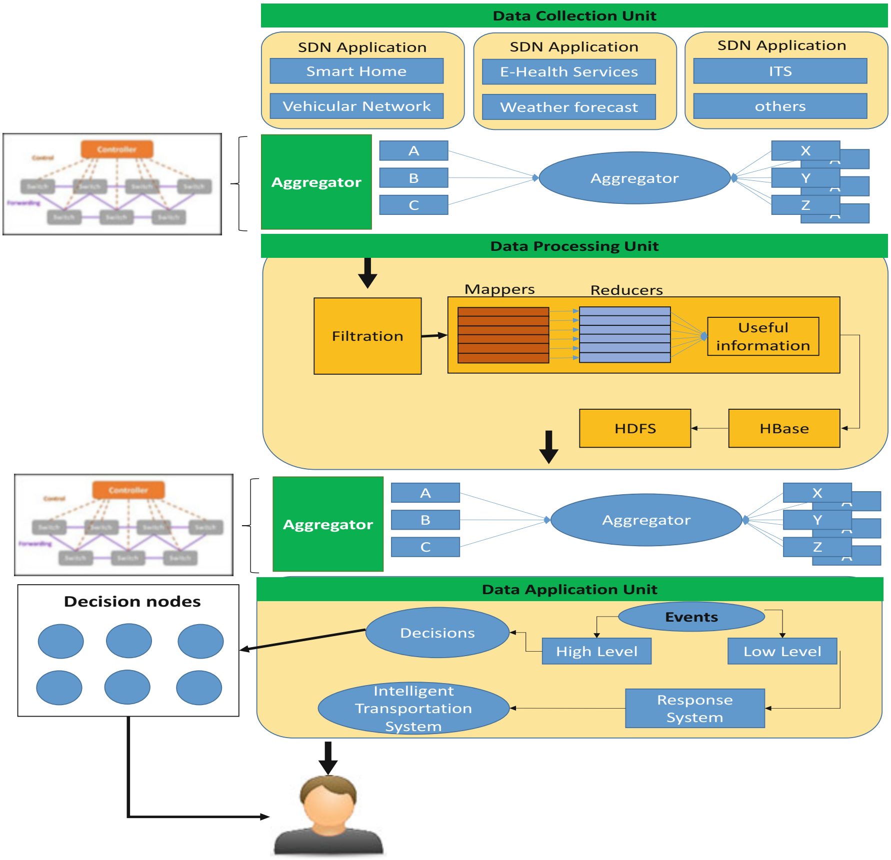


Fig. 6.1 Proposed internet of software-defined system architecture for smart cities

6.3.1 Proposed System Architecture

The method of execution flow in the proposed system is done in various levels, mainly three major levels and two intermediate support levels. Various embedded units or devices transmit data to the data processing level using standard networking technologies. During this communication, data is also received by the respective users who need it for their various jobs. Figure 6.1 shows the flow diagram of the proposed system architecture.

6.3.1.1 Data Collection and Intermediate Levels

In a smart city, few services are considered for data loading and collection purpose, for example, related medical data, public transportation, population, general facilities, etc. Normally sensors connect these services for the efficient data loading and collection. The data is communicated between layers through system ways, that is, data is passed to the upper levels using the first Intermediate level. The sensed data is transmitted through the aggregator points (AP) which are attached to sensors. To reduce the congestion, the APs are further classified into three levels, that is, zone, local, and global APs. These zone levels take the responsibility of loading high detail information from each unit in the smart city. The local level AP (LLAP) is the *n* charge for aggregating sensor data from similar units such as medical hospital of online health services, transportation data from roads, etc. Finally, the global level AP (GLAP) eventually collects data from the LLAP and send it to the SDN core network. The SDN core network uses the SDN controller to map the traffic received from the GLAP over the network. The controller is designed to perform many useful functions, for example, to differentiate sensed data based on the sensor IDs, topology control, optimizes the duty cycle of sensors attached with each AP, routing decision based on the user requirements. The SDN-enabled router using a priority application table to determine application-specific routing of data. In a typical scenario, under the IoT, many thousands of devices are interconnected for various services to produce high volume data. It is certain that because of the congestion, the network operation will slow down. The literature shows a number of methods developed by different researchers to identify congestion level on any given link. For example, Kandula in [20] developed a method based on a predefined threshold. It is suggested and proposed that up to 70% capacity used link can be declared as a congested link. A similar predefined threshold with 75% of link usage approach is adopted in this research. In order to efficiently transfer big data over SDN, the design should mention a constraint of this work. Therefore, the current traffic engineering technique available in the literature has been adopted to enhance the performance of the SDN. An outstanding suggestion will be given to use/choose a traffic engineering mechanism for controlling and routing traffic over an SDN. One example would be to use high-speed processing switches and routers because of the high volume data generated through the sensors in an IoT environment. The poll- and push-based methods are used by SDN controller to obtain useful information about the link (as mentioned above) statistics from switches. It is generally believed that poll based is a slower method as compared to the push-based mechanisms. Because of this feature, push-based method can be used for high-speed data routing and switching. Chen in [21] proposed double-level congestion handling schemes which are useful to be implemented. Chen in his work proposed that the SDN controller maintains a global view of the whole network by getting information of links from each switch in the SDN. This information is further used to control the load on each link in the network.

Table 6.1 Data collection of specific block

FV	Time	SensorID$_1$	SensorID$_2$	…	SensorID$_n$
F_1	t_1	$X_1{}^{(1)}$	$X_2{}^{(1)}$	…	$X_n{}^{(1)}$
F_1	t_2	$X_1{}^{(2)}$	$X_2{}^{(2)}$	…	$X_n{}^{(2)}$
…	…	…	…	...	…
F_n	t_m	$X_1{}^{(m)}$	$X_2{}^{(m)}$	…	$X_n{}^{(m)}$

Algorithm 6.1: Data Collection

```
Begin:
        IF   Data: = Required Data
                    Aggregate ()
        ELSE
              Partition Dataset into Blocks BL and BR
              Aggregate BL
              Aggregate BR
              Merge BL and BR
           END IF
END
```

Table 6.1 shows the data aggregation in general, where data of the particular block is aggregated at different time intervals (e.g., t_0, t_1, … t_m) based on sensorIDs.

FV is a specific attribute called factor variable. The FV is used to represent a particular attribute of the data block by which the noise is detected and discarded during data filtration process. The data is in raw and not complied; therefore, in the first instance the data is sent to the preprocessing stage to the data management and processing level which are explained as follows:

6.3.1.2 Data Management and Processing Level

To convert data to a meaningful form, the data received from the intermediate level 1 is sent to data processing level for normalization purposes. The normalized data is useful for the user to make a decision such as a road congestion data is useful for the commutator to make a decision to reach the destination as quick as possible and using a short route. As mentioned earlier that high processing power would require processing high volume data. Conventional *computational* methods are always challenged while processing real-time data. The latest technology such as Hadoop ecosystem and GraphX and SPARK are deployed by the data processing level in this kind of situation (real-time data processing). Further other capabilities of the Hadoop system such as HDFS and heterogeneous cluster Hadoop system are also used for processing, storing, and manipulation tasks related to huge data. The current literature abounds with many techniques to assign job to the Hadoop cluster system. In real-time processing, a division of huge jobs to sub job will be required through a proper scheduling mechanism a map-reduce part. After loading sub job to map-reduce system, the system allows to alter the number of jobs easily (if required). Therefore, an adaptive (intelligent) job scheduling technique is used in this research to adjust the load on the map-reduce system dynamically. A job tracker is assigned to each job which uses two different parameters to switch the part of the job from the current state to another state such as from processing unit to memory or any other part of the system. The change of sub job from one state to another state is executed in real-time environment and total based on the amount of a load of a cluster. This is not a big challenge in Hadoop system environment. However, fixed job assignment does not produce optimize results heterogeneous Hadoop clusters systems.

The proposed system (scheduling mechanism) provides a solution to two main problems because of the fixed job assignment, that is, (1) a high-performance node remains in the idle state and (2) a low-performance mode always remains in the busy state. This trade-off between high- and low-performance mode makes the system unstable for heterogeneous Hadoop clusters. In the proposed system, our developed job scheduling strategy checks the load on each node at runtime. A node always automatically demands new jobs, if its current workload is dropped from the threshold value of 75% of its total link capacity. During each cycle, the proposed scheduler checks the existing load of each node and loads new jobs accordingly. Hence, it optimizes the total capacity of a node by incorporating the load parameter during a single turn. The output of the map-reduce system is also passed to HDFS system for storing and other required operations.

Rules Definition Various threshold values which are a threshold limit lalue (TLV) are defined and set for evaluation of the datasets after the completion of the preprocessing stage. It is understood that TLVs are set for every dataset, for example, fire detection function, water consumption, etc. Normally TLVs are specific numeric values to represent some function as discussed above, for instance, 90% pollution, etc. Along with the TLV, certain rules are also defined to make a decision when that value is met or surpassed.

Algorithm 6.2: Rules Definition using IF/THEN
```
Rules Definition
Begin:
        IF Traffic_Data > TLV
               Traffic Congestion ()
         IF Temp > TLV
               Fire_Detected
         // and so forth
END
```

Algorithm 6.2 is used to define rules where TLV is the specifically defined threshold limit value and Temp are the temperatures.

Message Queue After the definition of rules, the rules and data are stored systemically in the message queue (MQ). Using MQ mechanism offers some advantages that do not necessitate an instant reaction to carry on the processing. It also avoids delay and halts during the processing. The working mechanism of the MQ is very straightforward by receiving M message at time t; the MQ takes action accordingly. Figure 6.2 shows the handler H which is handling the overall process in this regard.

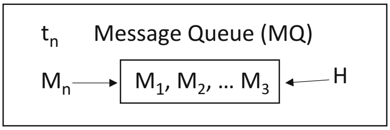


Fig. 6.2 Queue mechanism in SDN

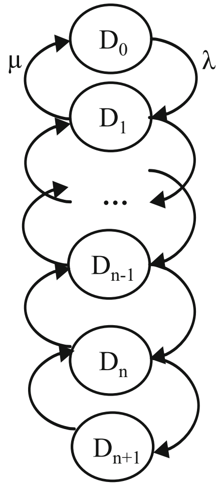
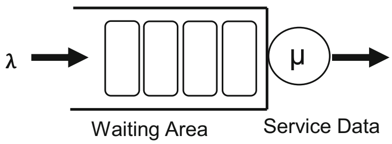


Fig. 6.3 M/M/1 queuing model for SDN

The proposed architecture has used basic M/M/1 queuing model because of its efficient functioning wherein M/M/1 model system (e.g., S system) is in a stable state if coming rate of the messages (data values) is less than the required service rate. This mechanism, which is used in the proposed architecture, is shown in Fig. 6.3 with the communication from one data item of the dataset to another. The mathematical process of the M/M/1 mechanism is illustrated as follows:

D_0 and D_1 are maintained balance with each other which is required during steady state, hence for $D_2, D_3, \ldots D_{n-1}, D_n, D_{n+1}$, such as:

$\lambda D_0 = \mu D_1$
$\lambda D_1 = \mu D_2$
$\lambda D_2 = \mu D_3$
$\ldots$
$\lambda D_{n-1} = \mu D_n$



After the abovementioned analysis, the following equation is obtained by rearranging the abovementioned equations:

$$D_n = \left(\frac{\lambda}{\mu}\right)^n D_{n-1} = \left(\frac{\lambda}{\mu}\right)^n D_0 \tag{6.1}$$

The overall probability which is known to be equal to 1, we get the following equations:

$$\sum\nolimits_{j=0}^{\infty} D_i = 1$$

$$D_0 \left(1 + \left(\frac{\lambda}{\mu} \right)^1 + \left(\frac{\lambda}{\mu} \right)^2 + \left(\frac{\lambda}{\mu} \right)^3 + \ldots \right) = 1$$

The abovementioned process and equation yield the sum of geometric series, therefore,

$$D_0 = 1 - \frac{\lambda}{\mu}$$

The utilization can be $D = 1 - D_0$, thus

$$D = 1 - \left(1 - \frac{\lambda}{\mu} \right)$$

$$D = \frac{\lambda}{\mu}$$

By applying values in Eq. (6.1), we get

$$D_n = \left(\frac{\lambda}{\mu} \right)^n D_0 \Rightarrow (D)^n \ (1 - D)$$

$$D_n = S^n \ (1 - D)$$

The system has j number of tasks to perform, then

$$M = \sum\nolimits_{j=0}^{\infty} j.D_j = \sum\nolimits_{j=0}^{\infty} j.D^n \ (1 - D) = (1 - D) \ D \frac{d}{dD} \sum\nolimits_{j=0}^{\infty} D^j$$

$$M = \quad (1 - D) \ D \frac{d}{dD} \left(\frac{1}{1 - D} \right) = \frac{1}{1 - D}$$

The average stay of job/task in a system is

$$T = \frac{N}{\lambda} = \frac{D}{1 - D} . \frac{1}{\lambda} = \frac{1}{\mu - \lambda}$$

Similarly, the average number of tasks in the queue is
Mq = M – tasks in services

$$\mathrm{Mq} = M - \mathrm{Tasks} = \frac{D}{1-D}.S = \frac{D^2}{1-D} = \frac{\lambda^2}{(\mu-\lambda)^{\mu}}$$

Waiting time in queue becomes

$$\mathrm{Wq} = T - \frac{1}{\mu} = \frac{1}{\mu-\lambda} - \frac{1}{\mu} = \frac{\lambda}{(\mu-\lambda)^{\mu}}$$

Similarly, the average number of jobs/tasks in the queue is

$$M = \frac{D}{1-D} = S + \frac{D^2}{1-D}$$

6.3.1.3 Application Level

The SDN receives data from the HDFS system at the application level via intermediate level 2 (IL2). There is no difference between the working of the IL2 and IL1 except for the level of the traffic is different on the SDN is less compared to IL1. The application level consists of two parts: (1) event and decision management and (2) named data network (NDN). In the event and decision module, an event is generated based on the data from the data and processing level. The result, that is, the event is communicated to the concerned departments to be processed for the respective user. The management module divides generated event into two groups, that is, high- and low-level events. As it is clear from the name that the high-level events are the most important events and processed on a priority basis where the other (low-level events) are saved in the decision module until a notification is sent back to the data processing level. After the proper acknowledgment and notification by the various required modules such as decision management module, it is processed and discarded. To fully understand the process of the event and decision module, it is explained through a comprehensive scenario.

Consider an example of a situation of the city where the city collects the data of a road congestion level. The data is communicated to the event and decision module. If the communicated data is more than the level of the predefined threshold which will be considered as a high-level event by the decision module and relevant messages are sent to the concerned department. On the other hand (e.g., the congestion level is less than the threshold value), the event is discarded after proper notification. During this process, all users will also be informed about the situation by the decision module as discussed earlier. Such transmission of data is performed to either fulfill user request or automatically broadcast to a group of users following the hierarchal model presented in Sect. 6.1. In both situations, ICN-based networks are preferred to send it to the user considering user's interest [22]. Therefore, a good way to use a named data network (NDN) to efficiently fulfill the user requirements either by sub/pub or pull-based communications as which is used in this research. Every decision module performs as an NDN node which consists of three entities, that is, pending interest table (PIT), content store (CS), and a forward information base (FIB). To avoid interest looping problem, PIT keeps the pending interests and their unique nonce values. In research presented in [23], it is performed using CS and FIB by storing incoming content and routing of content to another decision module. An interest packet sends it to the NDN whenever a user is interested in a particular data. The interested data is processed, and content is delivered from a decision module following the NDN.

6.4 Simulation Results, Implementation, and Analysis

As mentioned earlier that Spark and GraphX with Hadoop single node setup on UBUNTU 14.04 LTS coreTMi5 machine with 3.2 GHz processor and 4 GB memory has been used to implement the proposed system. Pcap packets are generated from the datasets by using Wireshark libraries and retransmit them towards the developed system. Hadoop-pcap-lib, Hadoop-pcap-side, and Hadoop Pcap Input libraries have been considered for network packets processing and generating Hadoop Readable format collection and aggregation unit so it can be processed easily by these systems (Hadoop and GraphX). To make smart transportation decisions, GraphX is used. The data has been taken from [9, 10, 24]. Varieties in the traffic intensities at various timings on the same road and the intensity analysis at various times of the day provide great support to the authorities to manage, plan, and monitor traffic on any particular time.

The intensity of the traffic of one of the roads of Aarhus city is shown in Fig. 6.4. It indicates that the morning hours between 7:00 and 9:00 and noon timing from 11:25 to 12:30, the traffic is heavy on the roads. The heavy traffic may be due to the school start and end timing and office start timing. It is a good performance by the proposed system which announces such a situation and issue alerts. It not only announces the heavy traffic but also looks into the blockage at roads if any. Various methods have been used to define the blockage which includes the number of vehicles and the average speed. If the average speed is less than a specific value (threshold), then it can be predicted that there is some sort of blockage at the road. The same information is communicated to the concerned users where they will try to find the time at which there is very less traffic.

As an example, the analysis is performed on Aarhus city traffic, and the analysis of the speed on the intensity of traffic is shown in Fig. 6.5. When the intensity of traffic is more, that is, more vehicles on the road between two points, the average speed of the vehicles is greater. The fall in some vehicles on the road results in a rise in the average speed. It can be easily noticed that in a high number of cars (in between 25 and 30), the average speed is very low at various times of the day, shown

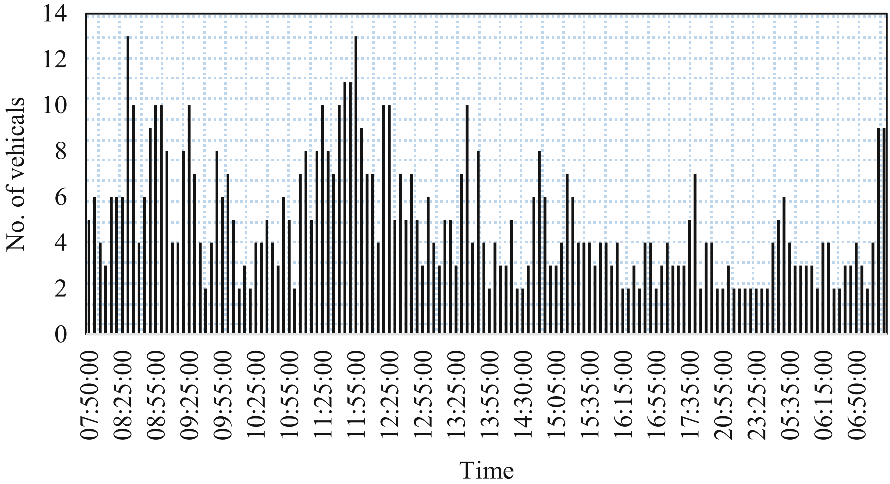


Fig. 6.4 Number of vehicles in a given time

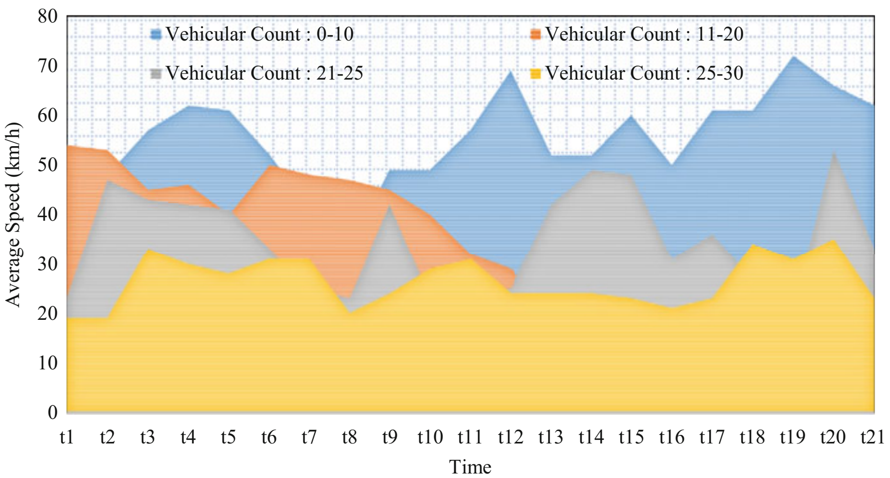


Fig. 6.5 The average speed of a vehicle

as a red color graph. In case of the low intensity in a number of cars (from 0 to 10) shown as a blue color, the average speed of the vehicles is quite higher.

The data shows some abnormalities with the less number of cars and other vehicles, and these abnormalities happen because of other reasons such as road is under construction or traffic is diverted to some road because of the incident. Usually, the distance is conserved to measure the time to reach the destination with an observation that the number of cars and speed (average) also affects the total time required to reach the target destination.

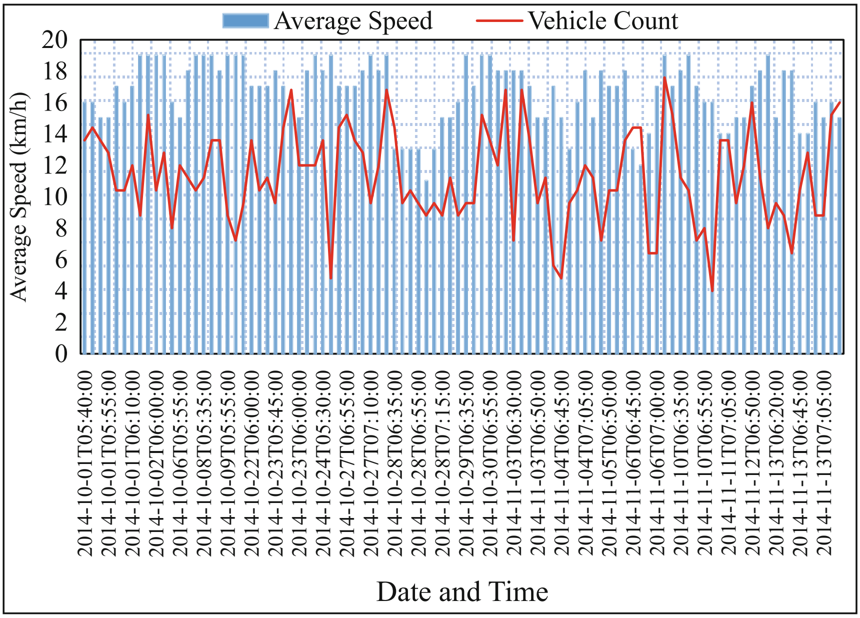


Fig. 6.6 Average speed during various dates and times

Figure 6.6 shows the road blockage in the city. Based on the proposed scheme, the average speed of the vehicles is too low even when there are a minimum number of vehicles on the road. It can be seen from various graphs that the road blockage is shown in the morning hours on different days. One of the main reasons for this is the construction work on the roads in the morning hours. In such cases, one can take real-time traffic information to calculate the shortest and quickest path between source and destination rather than only the distance information.

The humidity at home, shown in percentage in Fig. 6.7, plays a vital role in user behavior while doing some physical exercise or work. Furthermore, increase in the humidity demands more resources such as the usage of electricity, etc. The proposed system exploits this phenomenon through a learning mechanism where the sensor devices transferred data to the proposed scheme for experiencing the level of humidity. The proposed scheme in December (2016) takes into considerations some readings from the sensors and made threshold value. Using the past knowledge, the proposed system starts predicting the humidity level for January (2016) and others.

The prediction of humidity levels provides great support to users to decide once they know the level of the humidity as shown in Fig. 6.7. Figure 6.8 shows a similar data but outside temperature. The processing of large graphs to achieve smart transportation is the main contribution of this research work; the system is evaluated on efficiency regarding throughput (in megabytes/sec Mbps) and the response time (in milliseconds).

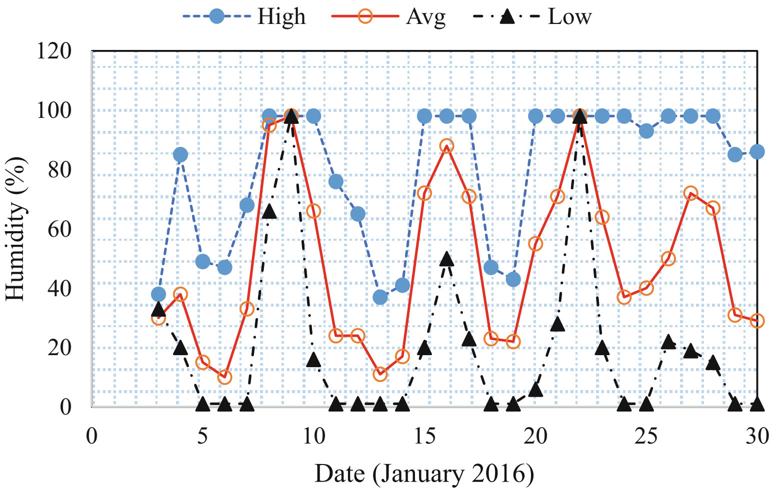


Fig. 6.7 Humidity inside home

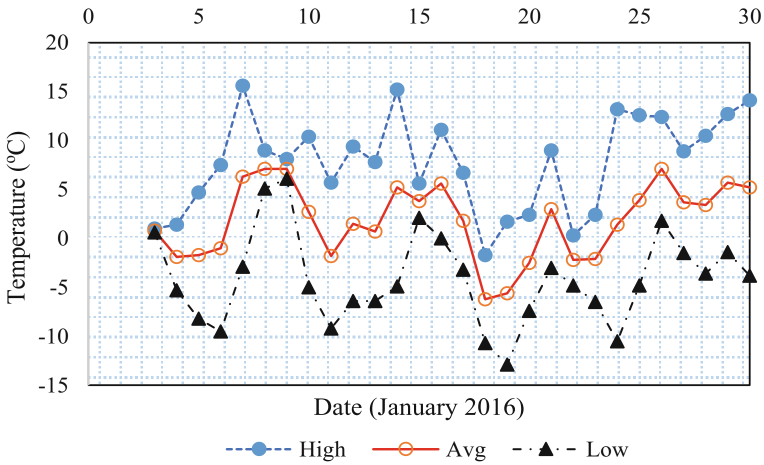


Fig. 6.8 Outdoor temperature

6.5 Conclusion

In this research, a novel technique is presented to analyze and process big data for useful application in a smart city environment. The latest technologies such as Hadoop ecosystem alongside Spark and GraphX are deployed while developing the abovementioned system. The novel technique is based on a three-level architecture for efficient data gathering in smart city environment using sensor-based devices. To normalize data for good use, a preprocessing stage is also used where the collected data is passed to the upper levels for processing using SDN and NDN. APs are used to get huge data and pass it to Hadoop ecosystem. A novel scheduling mechanism is implanted in the Hadoop ecosystem to balance the load on the Hadoop ecosystem efficiently. A comprehensive decision module is used, which performs and makes different decisions using the predefined threshold values. The user can also request in real time in the smart city environment to get support to make a decision.

A simulation test has been implemented for the proposed system on various data obtained from authentic sources. The simulation results indicate that the proposed system produced encouraging results. The use of Hadoop ecosystem was also outstanding and significant by analyzing the data with better speed and high throughput. Implanting cluster-based Hadoop system with sophisticated scheduling mechanisms to further reduce the processing will be one of the major tasks for future of this proposed research.

Acknowledgment This study was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2017R1C1B5017464).

References

1. IBM, Armonk, NY, USA. Four vendor views on big data and big data analytics: IBM [Online], 2012. Available: http://www-01.ibm.com/software/in/data/bigdata/
2. V.N. Gudivada, R. Baeza-Yates, V.V. Raghavan, Big data: promises and problems. Computer **48**(3), 20–23 (2015)
3. C. V. N. Index, Forecast and methodology, 2013–2018, 2013
4. N. McKeown et al., OpenFlow: enabling innovation in campus networks. ACM SIGCOMM Comput. Commun. Rev. **38**(2), 69–74 (2008)
5. N. McKeown, Software-defined networking. INFOCOM Keynote Speech **17**(2), 30–32 (2009)
6. K. Xu et al., Toward a practical reconfigurable router: a software component development approach. IEEE Netw. **28**(5), 74–80 (2014)
7. F. Al-Turjman, Fog-based caching in software-defined information-centric networks. Comput Electr Eng J **69**(1), 54–67 (2018)
8. Y. Han et al., Software defined networking-based traffic engineering for data center networks, in *Proc. 16th Asia-Pacific Network Operations and Management Symposium*, Taiwan, September 2014
9. S. Kolozali, M. Bermudez-Edo, D. Puschmann, F. Ganz, P. Barnaghi, A knowledge-based approach for real-time IoT data stream annotation and processing, in *Proc. of the 2014 IEEE International Conference on Internet of Things (iThings 2014)*, Taipei, Taiwan, September 2014
10. M. Gramaglia, O. Trullols-Cruces, D. Naboulsi, M. Fiore, M. Calderon, Vehicular networks on two Madrid highways, in *2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON),* Singapore, 3 July 2014, pp. 423–431
11. A. Ahmad, A. Paul, M. Rathore, H. Chang, An efficient multidimensional big data fusion approach in machine-to-machine communication. ACM Trans. Embed. Comput. Syst. (TECS) **15**(2), 39 (2016)
12. J. Rowley, The wisdom hierarchy: representations of the DIKW hierarchy. J. Inf. Sci. **33**(2), 163–180 (2007)

13. S. Bischof, A. Karapantelakis, C.S. Nechifor, A. Sheth, A. Mileo, P. Barnaghi, Semantic modeling of smart city data. Position Paper in W3C Workshop on the Web of Things: Enablers and Services for an Open Web of Devices, Berlin, 25–26 June 2014
14. R. Tönjes, P. Barnaghi, M. Ali, A. Mileo, M. Hauswirth, F. Ganz, S. Ganea, B. Kjærgaard, D. Kuemper, S. Nechifor, D. Puiu, A. Sheth, V. Tsiatsis, L. Vestergaard, Real time IoT stream processing and large-scale data analytics for smart city applications, in *Poster session, European Conference on Networks and Communications 2014,* 2014
15. S. Din, A. Ahmad, A. Paul, S. Rho, Service orchestration of optimizing continuous features in industrial surveillance using big data based fog-enabled internet of things. IEEE Access **PP**(99) (2018)
16. A. Awais, M. Babar, S. Din, S. Khalid, M.M. Rathore, A. Paul, A. Reddy, N. Allah, Socio-cyber network: the potential of cyber-physical system to define human behaviors using big data analytics. Futur. Gener. Comput. Syst. (2018)
17. A. Awais, M. Khan, A. Paul, S. Din, M.M. Rathore, G. Jeon, G.S. Chio, Towards modeling and optimization of features selection in Big Data based social internet of things. Futur Gener. Comput. Syst. **82**, 715–726 (2018)
18. S. Din, A. Paul, A. Ahmad, J. Hong, Energy efficient topology management scheme based on clustering technique for software defined wireless sensor network. Peer-to-Peer Netw. Appl., 1–9 (2017)
19. S. Din, A. Ahmad, A. Paul, A cluster-based data fusion technique to analyze Big Data in wireless multi-sensor system. IEEE Access **PP**(99), 1–1 (2017). https://doi.org/10.1109/ACCESS.2017. 2679207
20. S. Kandula, et al., The nature of data center traffic: measurements & analysis, in *Proc. 9th ACM SIGCOMM Conf. Internet Measurement*, 2009, pp. 202–208
21. M.-H. Chen et al., A low-latency two-tier measurement and control platform for commodity SDN. IEEE Commun Mag **54**(9), 202–208 (2016)
22. S.H. Ahmed, S.H. Bouk, D. Kim, RUFS: RobUst forwarder selection in vehicular content-centric networks. IEEE Commun Lett **19**(9), 1616–1619 (2015)
23. S.H. Ahmed, S.H. Bouk, M.A. Yaqub, D. Kim, H. Song, J. Lloret, CODIE: controlled data and interest evaluation in vehicular named data networks. IEEE Trans. Veh. Technol. **65**(6), 3954–3963 (2016)
24. S. Uppoor, O. Trullols-Cruces, M. Fiore, J.M. Barcelo-Ordinas, Generation and analysis of a large-scale urban vehicular mobility dataset. IEEE Trans. Mob. Comput. **13**(5) (2014)

Chapter 7
Taking Cloud Computing to the Extreme Edge: A Review of Mist Computing for Smart Cities and Industry 4.0 in Africa

Eustace M. Dogo, Abdulazeez Femi Salami, Clinton O. Aigbavboa, and Thembinkosi Nkonyana

7.1 Introduction

Cloud computing concept has been built with the assumptions of good Internet connectivity, adequate bandwidth and low latency. But with the proliferation of Internet of Things (IoT) of resource-constrained smart devices, stringent Internet connectivity demand, high bandwidth, low latency, lower energy consumption, context-awareness, mobility and enhanced security requirements placed on applications and services [2, 39] in Smart Cities and Industry 4.0 concepts, as well as the huge amount of traffic and data expected to be generated, the traditional cloud-centric architectural arrangement no longer holds due to these assumptions. Cloud computing is therefore gradually evolving into new complementary concepts as edge and fog computing and also dew and mist computing to address the concerns mentioned above by extending the capabilities of cloud computing to the extreme edge of the network closer to the data generation source. Mist computing extends the concept of fog computing to the extreme edge of the network at the level of microcontroller and embedded nodes and is based on IoT concept; on the

E. M. Dogo (✉) · T. Nkonyana
Department of Electrical and Electronics Engineering Science, Institute for Intelligent Systems, University of Johannesburg, South Africa,
e-mail: eustaced@uj.ac.za; tnnkonyana@uj.ac.za

A. F. Salami
Computer Engineering Department, Faculty of Engineering & Technology, University of Ilorin, Nigeria
e-mail: salami.af@unilorin.edu.ng

C. O. Aigbavboa
Department of Construction Management and Quantity Survey, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa
e-mail: caigbavboa@uj.ac.za

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_7

other hand, dew computing is based on the Internet and client–server concepts, whereby on-premises computer provides functionality that is independent of cloud services, but also collaborative with cloud services [59]. The overall objective of dew computing is to enable access to cloud computing services in the absence of Internet connectivity.

In this chapter, an overview of mist computing model/architecture is provided together with the scope and evolving nature of the topic, its characteristics and comparison to cloud and fog computing, application scenarios, possible challenges and future direction. There exist several surveys on fog, dew and mist computing, but this chapter, therefore, does not intend to revisit these concepts in their entirety. However, this chapter aims to discuss mist computing in Smart Cities and Industry 4.0 in the context of Africa. The study is an extension of the work done in [19], which focused on how Nigeria and other developing ICT economies in Africa can benefit from cloud computing and its evolving and complementary implementations, challenges, drivers supporting its growth and future vision in the region.

Section 7.2 of this chapter provides overview on the overlapping and complementary key background concepts, namely, IoT, big data, cloud computing, edge computing and its implementation. The architecture of mist computing, its characteristics and its similarity and dissimilarity with cloud computing, fog computing and edge computing is covered in Section 7.3. In Section 7.4, the potential, application and use case scenarios of mist computing with respect to two emerging concepts, namely, Smart Cities and Industry 4.0 with focus on Africa, are discussed. In Sect. 7.5, an outline of the possible challenges in the implementation of mist computing, based on observation and trends, is discussed together with the drivers supporting the growth and technical recommendations in the context of Africa. Section 7.5 also discusses the future perspective of Smart Cities and Industry 4.0, while Sect. 7.6 concludes this chapter.

Several wide-range and inclusive surveys [2, 8, 17] and studies have been undertaken in recent times on cloud computing [13, 45, 56], edge computing [2], fog computing [10, 25] and its implementation in Smart Cities [38] and Industry 4.0 [33], mobile edge computing [1, 29, 30, 53], cloudlet [37, 48, 49], dew computing [44, 52, 59, 60] and mist computing [39, 40].

The web search interest over a period of 12 months as measured by Google search trends on fog, edge, mist and dew computing is shown in Fig. 7.1. The Google trend search shows a very low interest and possibly low awareness overall on these evolving cloud computing concepts, most especially in Africa. This motivates the essence and purpose of the research work presented in this chapter.

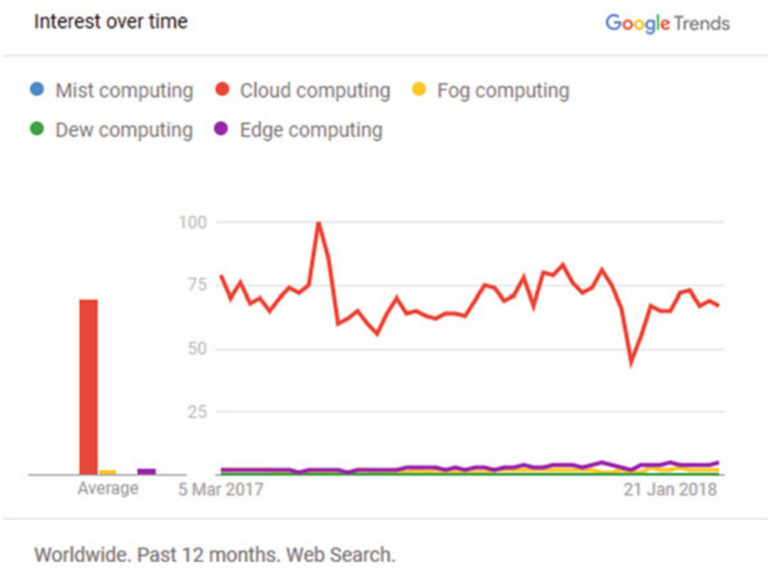


Fig. 7.1 Google search trend over the past 12 months

7.2 Overview of Key Background Concepts

In this section, an overview of key enabling concepts and technologies that are highly intertwined and overlapping with mist and dew computing is given as it relates to Smart Cities and Industry 4.0.

7.2.1 Internet of Things (IoT)

IoT is a communication technology concept that envisions a scenario where objects or things are interconnected with one another, based on standard communication protocols, each with its uniquely identifiable addressing system, forming an integral part of the Internet [6]. IETF categorizes IoT into three: people, machines and information. Figure 7.2 depicts the three-tier IoT functional stack architecture, comprising of the *Things* embedded with sensing capabilities, the *networking* layer and the specific *application* layer. IoT is vast and overlaps with other research fields as depicted in Fig. 7.3.

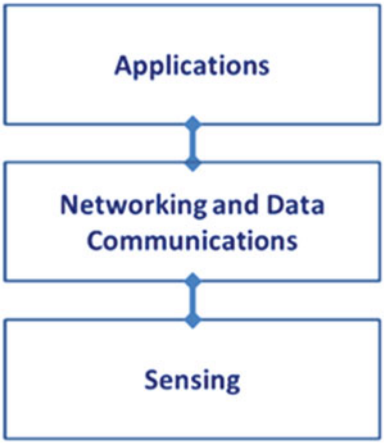


Fig. 7.2 Three-tier Architecture of IoT [31]

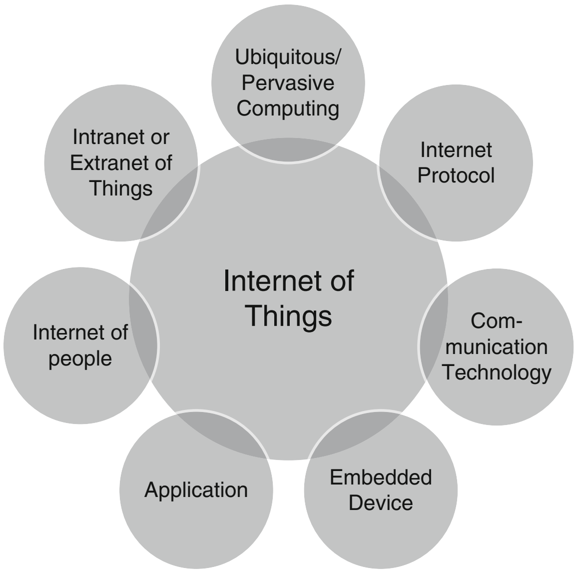


Fig. 7.3 IoT overlap with other research areas [31]

Wireless communication technology plays an important role in IoT, and [11] identifies the key and important aspects of IoT technology such as Radio Frequency Identification (RFID), Wireless sensor networks (WSN), Network Addressing and IoT middleware software. The authors also proposed a new paradigm called CloudIoT, which is the integration of cloud computing and IoT as these two complementary technologies are expected to shape the direction of the current and future Internet. McKinsey & Company predicts a potential economic value of up to \$11 trillion by the year 2025 if the policy makers and businesses can link the cyber-physical systems together. On the other hand, IDC predicts that by 2019, 45% of

IoT-created data will be stored, processed, analysed and acted upon close to/at the edge of the network.

The annual IoT Forum Africa (IoTFA), a platform that enables industry and expert practitioners to share their vision on IoT trends, challenges and solutions in virtually all aspects of life and in different industries, reinforces how IoT has the potential of transforming the African continent.

7.2.2 Big Data

The rate of deployment and expansion of devices and sensors connected to the IoT is a major source of data generated and transmitted in real time. Cisco projects that by 2020, the number of connected devices will exceed 50 billion [21], while some experts have even projected higher numbers due to the rapid advancement in the Internet and Internet device technology. It must be mentioned that with respect to these generated data, cloud computing facilitates storage, processing and analysis. Overtime data collection and analytics have evolved into three (3) ways – Analytic 1.0, collection of historical data; Analytic 2.0, social media and unstructured information; and Analytic 3.0, real-time IoT data from a vast number of sensors (heterogeneity) – making it challenging to push data to a single cloud data warehouse, hence the need to bring data analytics closer to the edge of network into routers and gateways, as well as on-board embedded systems with sensors (mist computing).

The characteristics of big data are captured in the 5 V's model, namely, volume, variety, velocity, veracity and value [35]; hence, a data that meets these characteristics is termed big data. Big data in Africa has numerous applications, such as in climate change, poverty and disease surveillance, agriculture, banking and finance, supply chain, media, space research and biological research [34, 41, 54].

7.2.3 Cloud Computing

According to the US National Institute of Standards and Technology (NIST) and the European Network and Information Security Agency (ENISA), "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to shared pool of configurable computing resource (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction, through visualization and multi-tenancy arrangement". Readers are referred to [19] for details on cloud computing concepts and analysis on how it relates to Nigeria and by extension Africa.

7.2.4 Edge Computing and Its Implementations

Edge computing (EC) has evolved from cloud computing due to the tremendous growth in IoT devices, which according to Cisco is estimated to reach 50 billion devices in no distant future. This comes with several challenges to the traditional cloud computing arrangement on several fronts such as the need for low latency, high bandwidth, enhanced security due to several organizations' sentiments on their data residing in unknown third party's data centre and the growing demand and advocacy for a greener technology. Hence, the need to extend cloud computing capabilities closer to the data source on the edge of the enterprise network using the edge gateway's core routers and switches is necessary. Overtime, edge computing has also evolved into other complementary implementations to address challenges and flaws associated with cloud and edge computing. A brief overview of the technological implementations of edge computing is discussed over the next subsections. For a more detailed survey, readers are referred to [8].

7.2.4.1 Fog Computing

Fog computing (FC) was originally proposed by Cisco [10], which extends cloud computing to the edge of the network at the level of routers and gateways. Its emergence is due to developing applications that are sensitive to latency, which cannot be met by the current cloud-centric architectural arrangement. There is currently no globally accepted definition of FC; however, for the purpose of the research work presented in this chapter, the definition formally stated by [10] is adopted, which states that: "Fog computing is a highly virtualized platform that provides compute, storage and networking services between end devices and traditional Cloud Computing data centres, typically, but not exclusively located at the edge of the network" with the fog nodes as its building blocks which are "distributed fog computing entities enabling the deployment of fog services and formed by at least one or more physical devices with processing and sensing capabilities (e.g., computer, mobile phone, smart edge device, car, temperature sensors, etc.). All physical devices of a fog node are connected by different network technologies (wired and wireless) and aggregated and abstracted to be viewed as one single logical entity, that is the fog node, able to seamlessly execute distributed services, as it were on a single device" [20].

Figure 7.4 depicts the fog to cloud architecture, with one cloud layer and two fog layers. Fog layer 1 is directly connected to the edge devices, while fog layer 2 is in between fog layer 1 and the cloud, acting as an intermediate processing layer. According to the OpenFog Consortium, advantages of fog computing include enhanced security due to the distributed architecture, cognition/decision-making on fog node, improved agility, lower latency and overall efficiency.

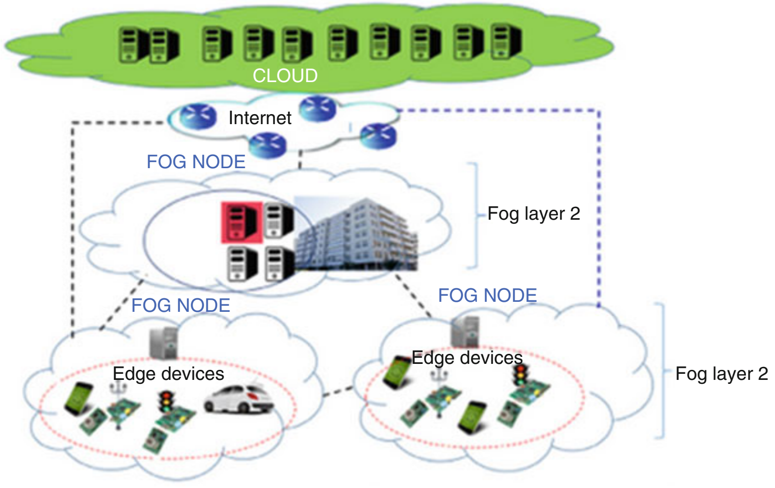


Fig. 7.4 Fog-cloud architecture [20]

7.2.4.2 Mobile Edge Computing

Mobile edge computing (MEC) is a telecommunication concept driven by smart mobile devices, IoT and 5G communications paradigms, which offers cloud computing capabilities and features in terms of storage and mobile computing at the edge of the radio access network (RAN) leveraging on the mobile base stations away from the traditional mobile cloud computing (MCC) arrangement, thereby reducing latency and enabling efficient use of the core mobile network and better mobile user experience [1, 30]. While IoT technology over the years is reaching its maturity, unlike 5G communication technology whose maturity is still being anticipated, there are already very high expectations due to the many potential advantages that next-generation 5G brings to wireless communication technology. Several expert opinions agree on the positive gains that Smart Cities-based [4, 32] and Industry 4.0-based [23, 47, 55] solutions can acquire by leveraging on future 5G network such as dealing with the exponential growth of data, user density, data capacity, low power consumption and data rates for very large connected IoT devices and catering for mission and safety-critical applications [18]. To this end, recently mobile edge computing (MEC) was renamed multiaccess edge computing to reflect the interests of both cellular and non-cellular operators across the industries [2, 53]. Ten (10) key technological components were identified and discussed in detail by [3] that will satisfy the requirement and building blocks of 5G wireless network communication technology, namely, a wireless software-defined network (WSdN), network function virtualization (NFV), millimetre wave transmission (mWT), massive MIMO, network ultra-densification (NUd), big data and mobile cloud computing (MCC), scalable Internet of Things, device-to-device connectivity with high mobility, green communications and new radio access network techniques (RAN).

7.2.4.3 Cloudlets

Cloudlet is the middle layer in a three-tier hierarchical architectural model made up of end-user mobile devices–cloudlet–cloud, similar to the arrangement in fog computing for IoT devices. This model's arrangement provides for an edge–cloud platform that extends cloud services and features closer to the end-user mobile devices [48, 49)]. Cloudlet can be viewed as the edge of the Internet providing computing, storage and networking capabilities to neighbouring mobile devices acting as a thin client to offload resource-intensive task and data caching, thereby improving execution speed and energy savings [37, 53]. Cloudlets are developed with virtualization features to meet the current demands of emerging mobile applications, such as high computational resource demand of mobile applications and service with low latency requirements.

7.2.4.4 Microdata Centre

Microdata centre (MDC) is a modular pre-fabricated standalone data centre unit enclosed with all the features of a traditional data centre such as servers, virtual machines (VM), cooling system, network connectivity, uninterruptible power supply (UPS), security and access control systems as well as flood and fire protection and a high degree of mobility and ease of deployment. Bandwidth constraints, growth in time-sensitive applications and security are some of the driving factors behind the adoption of MDCs by organizations as part of an edge cloud solution, in order to support IoT workloads at the edge of the enterprise network in a distributed manner. This is a viable solution not only for established organizations but also for SMEs and developing regions where real-time data analysis is required. MDC finds application across different industries such as assembly plants, manufacturing industries, exploration and mining, smart cities application solutions and construction industry [8].

7.2.5 Fluid Computing

Fluid computing is an envisioned next-generation cloud architectural paradigm that eliminates the segregation created by cloud, fog and mist computing technologies, unifying them under a single abstraction through provisioning of an end-to-end cooperative platform which can be used for seamless computing, storage and networking to allow application data flow between functionalities regardless of whether the resource has been provided by cloud, fog or mist infrastructure. This is done by leveraging on virtualization and soft computing [12, 16].

7.2.6 Blockchain Technology

Security, privacy and regulatory frameworks [5] are key concerns relating to new and emerging technologies in cloud computing, Smart Cities and Industry 4.0, due to the sensitive information generated and transmitted by resource-constrained smart terminal devices in the IoT ecosystem which are prone to security attacks. Blockchain technology, a peer-to-peer (P2P) distributed ledger technology for transparent transaction devoid of a trusted intermediary, has evolved beyond its original application to support bitcoin cryptocurrency and is seen as a viable option to address these concerns by integrating it with smart devices in smart cities in order to provide a secure communication platform [9] and with cyber-physical systems (CPS) in Industry 4.0 to provide transparent transactions among smart devices [7].

7.3 Extreme Edge Computing

7.3.1 Mist Computing

Mist computing has evolved from fog and cloud computing into a collaborative cloud technology based on the idea that communication should be made at the level of sensors and actuators without burden on the communication network and the Internet, by leveraging on the computational networking resources from the devices at the very edge of the IoT networks. This leads to increased autonomy of the system and reduced latency and bandwidth usage with a corresponding power consumption reduction as communication accounts for five times consumption compared to computation. “Mist Computing represents a paradigm in which edge network devices, that have predictable accessibility, provide their computational and communicative resources as services to their vicinity via Device-to-Device communication protocols. Requesters in Mist can distribute software processes to Mist service providers for execution” [28]. Scalability, reconfigurability, location self-awareness, situation awareness and attention and machine-to-machine communication (M2M) are core features of mist computing [39]. Cisco is leading the vision to move data analytics closer to where the data is generated for better decision support of mission critical services instead of backhauling it to the cloud or even the fog. IoT smart devices should not just be about data collection; most importantly, they should bring intelligence to the extreme edge of the network to derive value from the data for quicker decision support. Fog computing has been able to address a lot of challenges associated with IoT, including bandwidth and latency requirements. In fog computing, the gateway is key to ensuring the coordination and functioning of IoT applications; however, this arrangement has some drawbacks, such as the gateway being the single point of failure since the entire network is dependent on it [40]. For this reason, mist computing is gaining popularity fast and is seen to address some of the concerns on cloud and fog computing architectures with Thinnect, an IoT edge network service provider, at the forefront of the development and implementation of mist computing in real-life scenarios.

7.3.1.1 Architecture of Mist Computing

Mist is made up of a large number of heterogeneous devices at the extreme edge of the IoT network which is capable of providing some form of services to aid the improvement of IoT applications in terms of the computational processes [28]. Figure 7.5 depicts the conceptual framework of mist in the IoT environment, with the sensors and actuators as the generators and processors of data at the extreme edge of the IoT network layout.

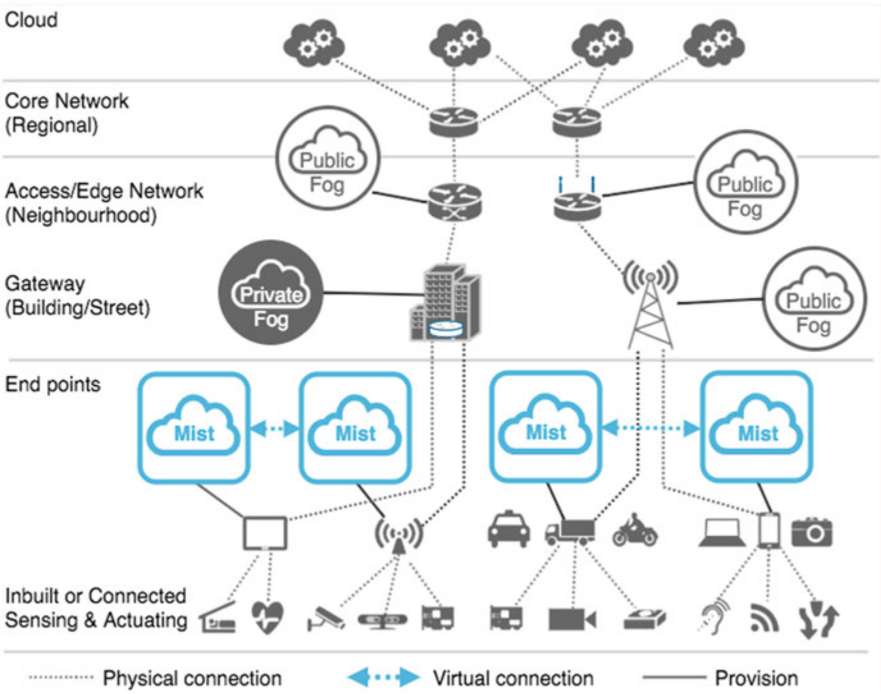


Fig. 7.5 Mist in IoT framework. (Source: Cisco)

7.3.1.2 Characteristics of Mist Computing

1. Hardware: Low latency, limited communication bandwidth, insufficient security, low mobility, low power
2. Situation awareness: Awareness on the physical environment and the local situation in collaboration with cloud and fog computing (see also Sect. 7.3.3)

7.3.2 Dew Computing

Since the emergence of the concept of dew computing originally proposed in [58], other technical definitions have also emerged from various researchers [22, 52]. In the context of the research work presented in this chapter, the definition in Wang (2016) [59] is adopted, which states that: "Dew computing is an on-premises computer software-hardware organization paradigm in the cloud computing environment where the on-premises computer provides functionality that is independent of cloud services and is also collaborative with cloud services". The goal of dew computing is to fully realize the potentials of on-premises computers and cloud services. The vision behind dew computing is based on the concept of the Internet, whereby it envisioned users situated in any part of an enterprise network to access the cloud without the continuous access to the Internet using a client–server model, in a complementary manner to cloud computing [46]. Figure 7.6 depicts a conceptualized model of dew–cloud architecture comprising of dew server, dew DBMS, dew client program, dew client service application and high-speed connection to the cloud server. For more details on the concepts of dew computing, readers are referred to [44, 59, 60].

7.3.3 Comparison of Mist with Cloud, Edge and Fog Computing

A summary of the comparison between cloud, fog and mist computing is shown in Table 7.1, which is based on earlier studies by [25].

7.4 Potentials, Applications and Use Case Scenarios of Mist Computing

In this section, the potentials, applications and use case scenarios of mist computing will be broadly discussed with respect to two emerging conceptual domains, namely, Smart Cities and Industry 4.0.

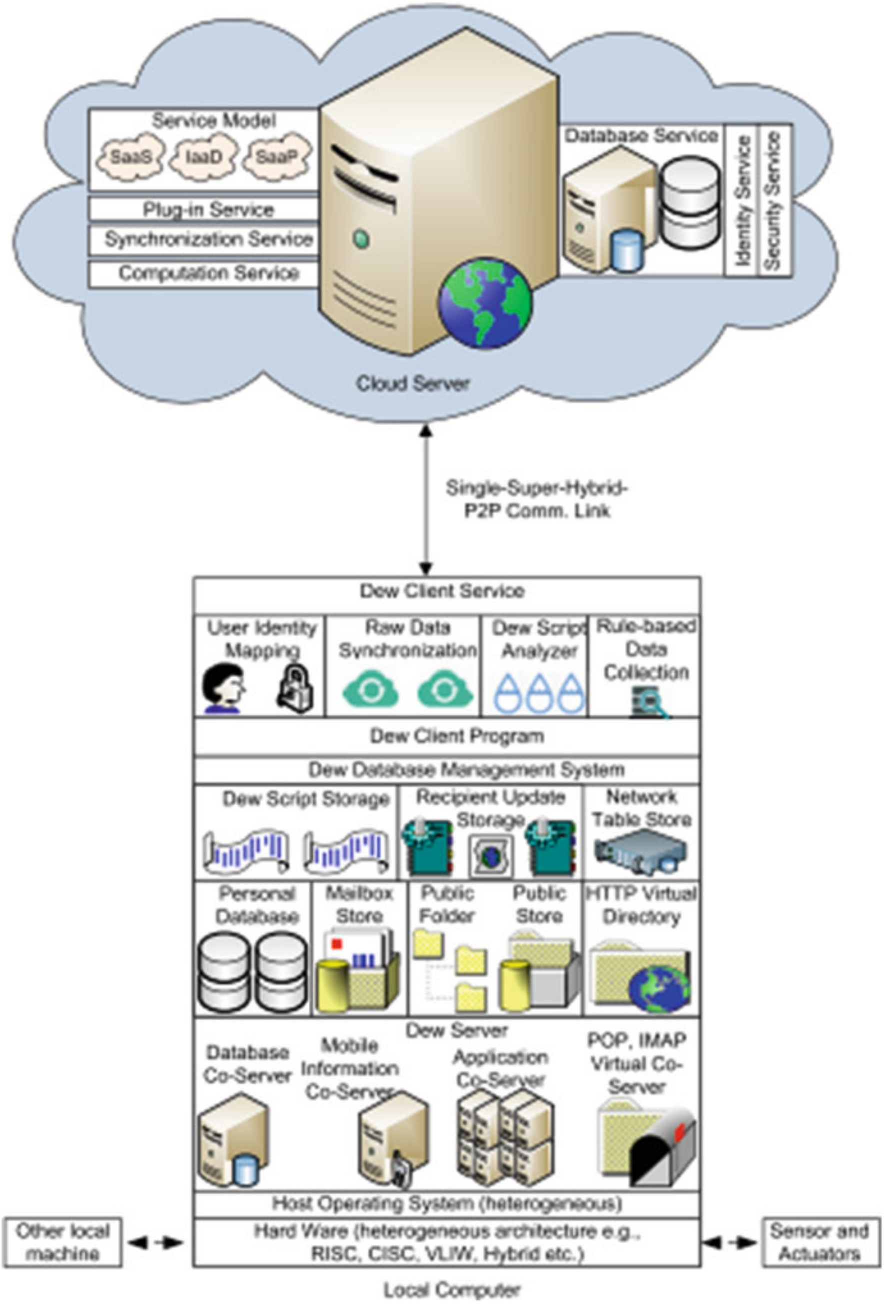


Fig. 7.6 Conceptual cloud–dew architecture [44]

Table 7.1 Comparison of mist with cloud, edge and fog computing

	Cloud computing	Edge computing	Fog computing	Mist computing
Support for IoT	Yes	Yes	Yes	Yes
Infrastructure deployment model	Centralized and accessed via the internet	Centralized	Decentralized LAN or IoT gateway/distributed	Centralized/distributed with the microcontroller network
Efficiency	High	Higher	Higher	Higher
Latency	High 100 ms and above	Low	Low 10–100 ms	Ultra-low 10–100 ms
Bandwidth	High 1Mbps–10Gbps	Low	Low 100 kbps–1Mbps	Low 1–300 kbps
Computational power	High	Moderate	Moderate	Low
Mobility support of IoT services	No	Yes	Yes	Yes
Data status	Data at rest	Data in transit/at rest	Data in transit/at rest	Data in transit
Security and privacy concerns	Higher concern due to sending raw data over the internet	High	Medium concern	Low concern since data reside with the hardware
QoS	Yes	Yes	Yes	Yes (limited)
Geographical availability	Limited spread in hundreds	Limited spread in hundreds to thousands	Wide spread in thousand to million	Very wide spread in million–trillion
Service type	Global	Limited	Limited	Very limited
Hardware	Large	Limited	Limited	Very limited
Working environment	Indoor data centre	Indoor	Outdoor/indoor	Outdoor/indoor
Power consumption	Very high	High	Low	Very low
Online/offline capabilities	Only online		Dual online/offline	Dual online/offline
Addressing	IPv4/IPv6	IPv4/IPv6	IPv4/IPv6	IPv6 (6LowPAN)
Cost of implementation	Medium	Medium	High	Low

7.4.1 Smart Cities

African cities are vital centres of trade, commerce, innovation and many other important activities. Recent statistics estimate that by 2030, over 60% of the African population would be dwelling in cities. In addition to this fact, these cities approximately contribute 70% of greenhouse gas emissions and energy utilization while occupying and accounting for only 5% of the continental land mass [14, 26].

Apart from these environmental trends, African cities are presently experiencing an increasingly huge demand for food, water, shelter and building materials, risk and disaster control strategies, waste management schemes and pollution control strategies [26].

African cities are therefore under constant and immense pressure to provide better standard of living, offer quality amenities and services, address social and environmental issues, foster economic growth and trade competitiveness, promote regional urbanization and industrial mechanization, attract investments and reduce costs [26]. These pressures are necessitating African cities to look into smart solutions in order to tackle the plethora of problems and challenges on the ground. Such solutions will enable African cities to transform into smart sustainable cities where ICT innovations and high-tech social amenities are used to provide improved standard of living, quality services delivery, seamless urbanized operations, economic self-reliance and growth trends that will match up with the environmental demands, social requirements and economic needs of the present and future generations.

Industrial experts and field specialists have shown that mist computing is more suitable for realizing smart city solutions where streets intelligently adapt to dynamic events, conditions, situations and changes in the city [15, 16, 39]. An interesting innovation which can possibly serve as a model for African cities to analyse, adapt and adopt is the smart street light control system developed by *Cityntel* [57]. This system is based on next-generation flat wireless mesh network (for direct application and device layer communications) and mist computing stack where controllers are equipped with data processing and embedded intelligence features. These smart controllers carry out network-based and/or device-specific decision-making operations and rely on wireless communication for data transmission and reception with sensors and detectors, such as movement and road surface detectors and noise, light reflection, air pollution, humidity and temperature sensors [57].

Unlike existing conventional systems, this solution replaces centralized permanent remote server control with distributed intelligent adaptive and autonomous situation- and context-aware devices. In this system, the LED street lights dynamically adjust brightness based on updates and notifications regarding the status of weather condition, human presence and movement and nature of traffic flow or intensity. In a situation where there is little traffic, street lights will switch to energy-saving mode by dimming their lights. Centralized computing is also utilized for processing aggregated data in order to make results of statistical analysis readily available for low-level device programming and learning. Energy conservation, minimized cost, easy upgrade and compatibility, fast and cost-effective deployment and high precision and reliability are ensured due to the adoption of adaptive, local communication and distributed computing strategies among low-cost wireless sensors for desirable system performance [57].

7.4.1.1 Smart Cities and Blockchain Technology

Despite all the numerous potential benefits associated with the concept of smart cities, information security and privacy remain a concern to stakeholders due to the vast amount of vital statistics and transit information generated by interconnected smart devices in the IoT ecosystem. There is, therefore, a need for information to: (i) be confidential, non-disclosure of sensitive information; (ii) possess integrity, information is accurate and reliable; and (iii) not manipulated or corrupted and available whenever/wherever needed by authorized persons. A lot of research has been carried out to address cloud security, but challenges remain due to the heterogeneous nature and resource constraints of smart devices, compatibility and other pertinent factors. A comprehensive review of cloud security is discussed in [27]. Recently, Blockchain which is a P2P distributed ledger technology devoid of a trustworthy intermediary (originally developed to support cryptocurrency) has received a lot of attention as a solution to address security concerns in numerous domains including smart cities [9]. Figure 7.7 shows a proposed smart city security framework across the physical, communication, database and application layers of IoT device to enable a common framework for secured data communication by different smart devices in a distributed community of devices.

7.4.2 Industry 4.0

Industry 4.0 can be conceptually viewed as embracing and infusing the core ideas of mist computing in the factory with the goal of gaining operational and performance efficiencies through rapid and precise decision-making for automated devices in the factory environment [24, 43]. In the context of this research, Industry 4.0 can be technically viewed as the collaborative use of mist computing, smart environmental sensors, robots, big data and mobile devices via cognitive and independent automation in order to achieve flexible mass production, easy product customization, scalability, predictability and greater efficiency during the entire manufacturing process [42].

This fourth industrial revolution goes beyond simply automating production, and it is based on cyber-physical systems achieved through complete and seamless digitalization of the manufacturing industry [24, 42]. Industrial locations and manufacturing operations are unique in Africa since the continental landscapes are vast with scattered and sporadic business hubs. In addition to this, many local corporations and manufacturing industries in Africa have matured to the level of making effective and productive use of data generating technologies and reactive data analysis strategies. It must also be added that a large number of manufacturing organizations in Africa are already aware of the ongoing global trends in Industry 4.0 together with the significant potentials and immense benefits of this explosive technology.

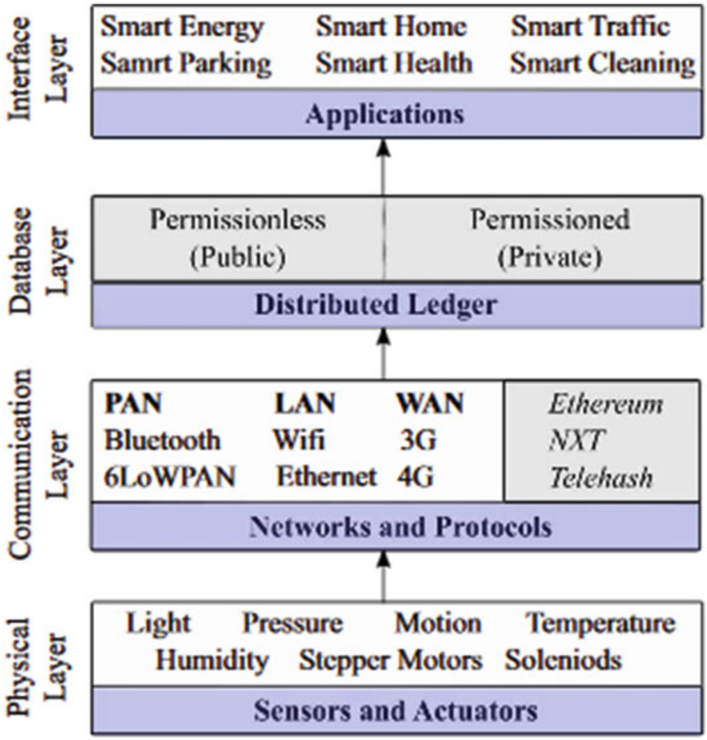


Fig. 7.7 Conceptualized smart cities security framework [9]

However, in comparison with developed ICT economies of the world, the African continent is still yet to attain full-blown maturity with respect to adopting, implementing and deriving maximum benefits from harnessing the potentials of Industry 4.0. Moreover, the widespread adoption of smart technologies and ICT innovations that can rapidly catalyse Industry 4.0 is still at a budding stage in most indigenous manufacturing industries in Africa. This impediment is a result of the general reluctance to venture into and invest in new knowledge and novel technologies by the government and industry of most countries in Africa. This hindrance is also due to the entrenched business culture among African manufacturers where cost-minimization measures are highly prioritized and expenditures on innovative technologies are curtailed to generate maximum profit in the prevailing harsh and competitive economic environment.

An exception to this is the case of African-based large international organizations with cross-continental operations where the level of adoption of smart solutions and innovative technologies is encouraging. From the African perspective, the proper adoption and judicious usage of Industry 4.0 offers limitless opportunities to tackle some socio-economic problems and prevalent industrial challenges associated with supply chain logistics through the development of smart machines, smart factories, smart plants, smart products, smart technologies and smart services. This will foster the development of unique, indigenous, high-tech gadgets, products, devices and services that can successfully leapfrog and dynamically compete with competitors in the global market. A suitable and relevant use case model for this research context is Siemens' *MindSphere*, which has the potential of serving as an underlying platform for mist computing services pertinent to resource optimization, energy data management and preventive/predictive maintenance [50, 51].

7.4.2.1 Industry 4.0 and Blockchain Technology

Blockchain technology which was originally developed for Bitcoin and cryptocurrency as aforementioned is also gradually finding interesting and fruitful applications in Industry 4.0. Supply chain management, smart contract, digital currencies and tighter cybersecurity are applications of Blockchain technology in Industry 4.0. The need for confidentiality, integrity and availability of sensor data is critical to industrial applications and manufacturing systems. Therefore, it is envisioned that Blockchain technology will catalyse remote machine diagnostics and machine-to-supplier transactions which will lead to improved spare parts replacement and overall maintenance practices.

7.4.2.2 Predictive Maintenance

Maintenance is a critical issue in manufacturing, and it goes without saying that maintenance is a prerequisite for effective and sustainable service delivery. Predictive maintenance (PdM) is one of the landmark innovations ensuing from Industry 4.0 which paves the way for predicting future asset failures, forecasting residual machine service lifetime, optimizing timing and scheduling of maintenance, unambiguously ascertaining machine operational statuses, proffering efficient preventive measures and enhancing equipment performance, quality and availability [36]. This is made possible due to the fact that historical and real-time data collected by the Industrial Internet of Things (IIoT) sensors is an integral part of the process of predicting when a failure is likely to occur. In this scenario, mist computing where sensors analyse data in real time close to the source is more viable than sending all the data to a cloud or fog gateway, as this saves time and quicker decision support in a more secure environment.

Benefits of PdM

Cost Reduction PdM with IIoT could benefit industry by improving efficiency, reducing inventory carrying costs, saving material costs, improving the return on asset (ROA), cutting down the cost of engaging service providers, reducing the cost of repairs and spare parts purchase and reducing unwanted costs associated with unplanned downtime and frequent maintenance. In addition to this, PdM with IIoT right on site reduces burden on network and overall cloud cost.

Direct and Impactful Value Creation PdM enables data collection without additional cost as the IIoT sensors are all located where the source of the problem is likely to occur. This makes it possible for prompt and rapid remote intervention in urgent repair and maintenance operations. Resultantly, manufacturing industries employing PdM achieve higher customer satisfaction and add more value and superior quality to the products and services offered to their customers due to their proactive decision-making processes, improved production strategies and pre-emptive business models.

Secure, Safer and Sustainable Operation By adopting PdM, enhanced data security is ensured as data is collected and analysed right in the factory instead of sending to the Internet or external data centre or gateways which are prone to cyberattacks. This enhancement enables manufacturers to deal with unexpected events and better reduce health, safety, environmental and quality risks in the entire manufacturing process. In addition to this, there is the advantage of easily and safely planning service intervals, achieving real-time analytics and attaining ultra-low response time during machine operation.

Uptime Improvement and Availability PdM ensures increased uptime and availability of the machine and equipment due to the reliance on automated intervention of autonomous and cognitive systems that ensure better machine availability and service life, prolong lifetime of aging assets, prevent lasting damage to relatively new equipment and foster quality control and precision manufacturing.

Accurate Predictions and Better Decisions By utilizing PdM, a plethora of new data sources are constantly made available that reliably and accurately reflect the real operational status of equipment for better decision-making. In addition to this, PdM ensures that valuable time is expended on data-driven problem solving and not wasted on brute-force data collection and validation. Another clear advantage of PdM in terms of economy and efficiency is that IIoT sensors, data collection, data analytics and decision support for PdM are all with a local area network (LAN).

7.5 Implementation Challenges in Africa

Skills Gap and Digital Divide The digital divide between developing ICT economies in Africa and developed ICT economies is not an overstatement, but this divide is still present and significantly wide in some regions of Africa. This digital divide, unfortunately, hinders or delays the rate of technical training, capacity building and skills acquisition necessary for the successful implementation of mist computing. In addition to this, the present engineering and ICT curriculum in most African institutions need reforms that will sufficiently enrich the programmes with practical skills that will adequately prepare and equip Africans for mist computing solutions and other emerging innovative technologies. Presently, there is very little research work that decisively investigates and quantifies the real impact of this existing digital divide and skills deficit. Therefore, it is necessary in Africa to embrace a holistic approach that considers the dual goal of investing in smart people and smart technologies.

Technology Transfer, Assimilation and Domestication It would be very challenging to realize the latent benefits of mist computing when solutions are simply imported from developed ICT economies into Africa as turnkey products and services. These smart technology solutions need to be meticulously studied, properly understood and, most importantly, technically adapted to be indigenously relevant, suitable and effective to tackle the specific development needs of Africa. This is necessary because economic situation, geographical context and cultural inclination play an indispensable role in the technology transfer process. Therefore, governments and corporate organizations in Africa need to adopt mist computing solutions in accordance with their prevailing regional socio-economic needs. For instance, adopting a mist computing solution that demands massive data aggregation from a large wireless sensor network for distributed intelligent transport and traffic management system will be prohibitively expensive and unsuitable based on the existing economic realities and social demands in most African countries. In such a scenario, a localized and simplified version of this smart technology solution that ingeniously harnesses the pervasiveness of mobile telephone data and other available cost-effective and easily accessible alternatives will be more suitable to the context of Africa.

Inclusiveness and Governance Smart technology projects involving mist computing call for smart governance models in order for such technological enterprise to thrive in Africa without suffering premature and sudden downfall. The need to develop and implement this pragmatic governance model in Africa is long overdue as there is an evident disconnect, disharmony and lack of synergy between polytechnics, universities, indigenous industries, multinational corporations, civil societies and local, state and federal governments. Such a smart governance model will bring together and carry along all parties during the course of executing the project in order to gain varying perspectives and get robust recommendations and vital suggestions from experts in the economic, legal, educational, political, technical, financial, ethical, social and developmental fields. The current governance model for technology projects in most African countries is still not fully inclusive due to information isolation which acts as a great obstacle to resource integration in the process of executing technological projects at both managerial and technical levels. Therefore, there is a need to reform the existing modus operandi and evolve smart governance models that will enable data and services from mist computing solutions to be useful for citizens, accessible for decision-making, readily available and highly effective in tackling the various developmental needs of Africa.

Failure, Faults and Risk Other challenges hindering the full adoption and implementation of smart technology projects involving mist computing in Africa are the failing infrastructures, resource constraints, high maintenance and running costs, faulty amenities, dysfunctional basic services, deleterious environmental conditions, counterfeit spare parts and unreliable products. These are part of the factors discouraging well-established global data players in mist computing from fully investing in most African countries. In addition to this, the extra cost of importing foreign quality products and services and the worrisomely slow rate of infrastructural development are major inhibitors for the full and permanent establishment of smart technology centres involving mist computing in most African countries.

Funding Opportunities and Business Blueprint Smart technology projects involving mist computing need financial resources and a high level of cooperation and partnership between the public and private sectors in order to ensure smooth running and completion of such projects. Unfortunately, respective governments of African countries are yet to come out with a clear, focused and robust business model and blueprint that will stipulate regulatory framework, infrastructural requirements, taxation guidelines, ownership rights and stakeholder policies in order to genuinely convince and encourage the private sector to invest and innovate in mist computing solutions and other related smart technology projects. Embarking on a smart technology project in Africa without this coherent and realistic business plan will unnecessarily complicate and delay the entire execution process which inadvertently opens the door for many risks and unforeseen setbacks. In addition to this, citizens are, most of the time, not fully and properly sensitized on the actual costs incurred and tangible socio-economic benefits ensuing from these smart technology projects. This makes it very difficult for the respective governments and corporate bodies in various African countries to enable the citizens to fully patronize and comply with the rules and charges associated with using such innovative technologies.

7.5.1 Drivers Supporting Growth in Africa

Trainable and Resourceful Young Population One of the biggest drivers encouraging the establishment of smart technology projects involving mist computing in Africa is the large youthful population who are able, hardworking and willing to learn and be trained on the technical and managerial aspects of these innovative technologies. These vibrant youths are very much cognizant of the global technological trends, and they desire to consume and enjoy the services of the same high-tech quality products as other consumers in developed ICT economies. Moreover, these African youths are relatively more entrepreneurial, dynamic, flexible, trendy, open-minded and computer literate, making it very easy and less costly for them to be trained on the proper understanding, usage, running, handling, operation and management of smart technology systems involving mist computing.

Limited Compatibility and Upgrade Bottlenecks Adopting smart technology projects involving mist computing will be relatively smoother, easier and quicker as issues tied to resource provisioning and the prohibitively high cost of upgrading and maintaining old technological infrastructures and outdated systems are more or less non-existent in the African context. This means that in the case of most African countries, there is the smart choice of starting with the latest and most up-to-date mist computing solutions available which automatically bring these African countries at par with well-established data players and competitors in the global market. Therefore, this makes it easier for African countries to successfully partake in the global market and reach a far wider audience and consumers with considerably less overhead costs and bottlenecks.

Industrialization Wave and Urbanization Drive Respective governments of various African countries are fully aware that industrialization and urbanization are keys to surviving and surmounting the onslaught of socio-economic challenges. Consequently, most African countries are now adopting a vigorous and aggressive innovation and digitalization drive that cuts across for-profit and non-profit organizations: educational, financial and government institutions and all other facets of life. With governments and citizens eager to improve their standard and quality of living, this powerhouse of innovation can significantly act as a catalyst for the widespread reception, successful adoption and effective implementation of mist computing solutions and other related emerging technologies.

Booming Connectivity The advent, upsurge and success of mobile telecommunication industries in Africa have undoubtedly made the impossible achievable and paved the way for the connection of people even in the remote regions of the continent. With the fastest mobile subscription penetration rate in the world, Africa is at a vantage point and strategically positioned to better collaborate with global data players and world technology leaders in developed ICT economies in order to fast-track the development of indigenous smart technology projects involving mist computing.

Industrial Focus and Research Interest It is worth mentioning and highly encouraging to highlight the heightened focus and growing body of research work on innovative technologies, smart solutions and mist computing from the industrial and academic sectors in Africa. This shows that there is a genuine concern to harness the immense potentials of mist computing in order to seriously tackle the plethora of developmental challenges on the ground and achieve a sustainable transformation of the computing landscape of Africa. In addition to this, these progressive efforts will open up more diverse opportunities in African countries for their citizens, local companies and government bodies to properly prepare and make adequate plans for the tangible, derivable benefits attached to the design and development of applications and systems based on mist computing. Therefore, a more favourable and highly receptive atmosphere is created which will support the adoption and sustainable development of mist computing in Africa.

7.5.2 Technical Recommendations

Reform Curriculums and Intensify STEM Education A large portion of innovative technologies and mist computing solutions rest on a solid foundation in science, technology, engineering and mathematics (STEM) education. Therefore, there is the urgent need for pragmatic curriculum reforms in Africa at preuniversity, university, postgraduate and vocational levels that will effectively infuse special skills in these fields needed for smart solutions. In addition to this, it must be mentioned that smart technology projects are multidisciplinary in nature, and they require multidisciplinary teams who can think and make decisions with multidisciplinary perspectives. Therefore, there is the need for relevant institutions in Africa to promote and provide opportunities for multidisciplinary learning and research.

Promote Local Innovation Culture Promoting a culture of technology domestication entails engaging a broad range of stakeholders and concerned parties who are educators, administrators, innovators, legislators, entrepreneurs, researchers and citizen members. This synergy is indispensable in order to effectively implement smart technology projects involving mist computing that will appropriately address the socio-economic needs of citizens of respective African countries. In addition to this, African countries need to create policy environments and devise judicious means of linking existing innovation infrastructures with novel mist computing solutions.

Provide Platforms for Inclusive Governance Respective governments of African countries need to urgently take advantage of the many innovative technological applications and social platforms in order to interactively and regularly engage their citizens and all stakeholders on the trends, recent efforts, progress and proper management of smart technologies and mist computing. This will go a long way in creating accountability, transparency and most importantly, a participatory and inclusive governance.

Establish Quality Control and Maintenance Centres In order to attain and maintain a high standard of smart city solutions and allied manufactured technology products in Africa, there is the pressing need to establish effective technology control and maintenance centres adequately equipped with the constitutional powers to monitor, assess and enforce quality for innovative technologies. The establishment of many of such centres in Africa will instil the culture of quality among various local science, technology and innovation (STI) communities, favourably attract foreign direct investments, improve the quality of mist computing services and, most especially, reduce the cost of adoption, implementation and domestication of mist computing in Africa.

Make Feasible Plans and Seek Concrete Funding Opportunities To achieve sustainable development of smart city solutions, respective governments of African countries should come out with feasible monetization policies, convincing cost recovery mechanisms and practical execution plans that will minimize wastage of resources and maximize efficiency gains. This approach will unambiguously and seriously convince interested local, regional and international investors to provide necessary financial support for smart technology projects involving mist computing in Africa.

7.6 Conclusion

Smart Cities and Industry 4.0 are still in their budding stage in Africa. But with the rapid pace of technological transformation with its envisioned disruption to businesses and society, it is critical for Africa to join, with the possibility of catching up technologically with the rest of the world. The current trend suggests a massive shift away from a centralized cloud architecture towards a distributed architecture with more computing and intelligent capabilities at the edge or extreme edge of the network closer to users. Mist computing has the potential to solve many of the unique socio-economic challenges currently facing Africa in different spheres of life in a collaborative way with the existing cloud and fog computing frameworks. This can be pragmatically achieved by leveraging on IoT solutions and Internet technologies such as smart transportation, smart grid, smart exploration and mining and smart healthcare. With the adoption of these smart solutions, African cities and industries will be adequately equipped with technologies to respond to perennial problems in normal and emergency situations such as industrial hazards, natural disasters, conflicts and insurgencies, scarcity and resource constraints, infrastructural failures and inadequacies, conventional and cybercrime occurrences and other forms of socio-economic challenges embattling the African region in a sustainable manner.

Acknowledgements This work was funded by the Institute for Intelligent Systems and the Department of Electrical and Electronic Science at the University of Johannesburg, South Africa.

References

1. N. Abbas, Y. Zhang, A. Taherkordi, T. Skeie, Mobile edge computing: a survey. IEEE Int. Things J. **5**(1), 450–465 (2018). https://doi.org/10.1109/JIOT.2017.2750180
2. Y. Ai, M. Peng, K. Zhang, Edge computing technologies for internet of things: a primer. Digit Commun. Netw. **4**(2), 77–86 (2018). https://doi.org/10.1016/j.dcan.2017.07.001
3. I.F. Akyildiz, S. Nie, S. Lin, M. Chandrasekaran, 5G roadmap: 10 key enabling technologies. Comput. Netw. **106**, 17–48 (2016). https://doi.org/10.1016/j.comnet.2016.06.010
4. M. Al Amine, K. Mathias, T. Dyer, Smart cities: how 5G can help municipalities become vibrant smart cities. Accenture. (2017), https://www.accenture.com/t20170222T202102__w__/us-en/_acnmedia/PDF-43/Accenture-5G-Municipalities-Become-Smart-Cities.pdf. Accessed 15 Mar 2018

5. Accenture Accelerating the Journey to Cloud (2017), https://www.accenture.com/t20170111T213812Z__w__/us-en/_acnmedia/Accenture/Designlogic/16-3360/documents/Accenture-2017-Top-10-Challenges-07-Cloud.pdf. Accessed 13 Apr 2018
6. L. Atzori, A. Iera, G. Morabito, The internet of things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)
7. A. Bahga, V.K. Madisetti, Blockchain platform for industrial internet of things. J. Softw. Eng. Appl. **9**(10), 533–546 (2016)
8. K. Bilal, O. Khalid, A. Erbad, S.U. Khan, Potentials, trends, and prospects in edge technologies: fog, cloudlet, mobile edge, and micro data centers. Comput. Netw. **130**, 94–120 (2018)
9. K. Biswas, V. Muthukkumarasamy, Securing smart cities using blockchain technology. 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), Sydney, NSW, Australia, pp. 1392–1393 (2016). https://doi.org/10.1109/HPCC-SmartCity-DSS.2016.0198
10. F. Al-Turjman, Fog-based caching in software-defined information-centric networks. Comput Electr Eng J **69**(1), 54–67 (2018)
11. A. Botta, W. de Donato, V. Persico, A. Pescapé, Integration of cloud computing and internet of things: a survey. Futur. Gener. Comput. Syst. **56**, 684–700 (2016). https://doi.org/10.1016/j.future.2015.09.021
12. D. Bourges-Waldegg, Y. Duponchel, M. Graf, M. Moser, The fluid computing middleware: bringing application fluidity to the mobile internet. *The 2005 Symposium on Applications and the Internet*, Trento, Italy, pp. 54–63 (2005). https://doi.org/10.1109/SAINT.2005.63
13. R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, I. Brandic, Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Futur. Gener. Comput. Syst. **25**(6), 599–616 (2009). https://doi.org/10.1016/j.future.2008.12.001
14. J.E. Cham, Internet of things: decoding the IoT ecosystem. RIPE Network Coordination Centre (NCC), RIPE75 (2017), https://ripe75.ripe.net/presentations/36-Decoding-the-IoT-Ecosystem-RIPE75.pdf. Accessed 16 Mar 2018
15. A. Corsaro, Cloudy, foggy and misty internet of things. In *Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering (ICPE '16)*. ACM, New York, pp. 261–261 (2016). https://doi.org/10.1145/2851553.2858661
16. A. Corsaro, Fluid computing: unifying cloud, fog, and mist computing. Embedded Computing Design (2016), http://www.embedded-computing.com/embedded-computing-design/fluid-computing-unifying-cloud-fog-and-mist-computing. Accessed 13 Mar 2018
17. K. Dolui, S.K. Datta Comparison of edge computing implementations: fog computing, cloudlet and mobile edge computing. 2017 Global Internet of Things Summit (GIoTS), Geneva, pp. 1–6 (2017). https://doi.org/10.1109/GIOTS.2017.8016213
18. ETSI Industry Specification Group, 5G (2018), https://www.etsi.org/technologies-clusters/technologies/5g. Accessed 12 Mar 2018
19. E.M. Dogo, A. Salami, S. Salman, Feasibility analysis of critical factors affecting cloud computing in Nigeria. Int. J. Cloud Comput.. Ser. Sci. (IJ-CLOSER) **2**(4), 276–287 (2013). ISSN: 2089-3337
20. E. Marín-Tordera, X. Masip-Bruin, J. García-Almiñana, A. Jukan, G.-J. Ren, J. Zhu, Do we all really know what a fog node is? current trends towards an open definition. Comput. Commun. **109**, 117–130 (2017). https://doi.org/10.1016/j.comcom.2017.05.013
21. D. Evans, The internet of things: how the next evolution of the internet is changing everything. Cisco Internet Business Solutions Group (IBSG) (2011), https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf. Accessed 13 Mar 2018
22. D.E. Fisher, S. Yang, Doing more with the dew: a new approach to cloud-dew architecture. Open J. Cloud Comput. (OJCC) **3**(1) (2016). ISSN 2199-1987
23. W. Harding, 5G to enhance industrial IoT and transform industry starting in 2020 (2017), https://industrial-iot.com/2017/04/5g-enhance-industrial-iot-transform-industry-starting-2020/. Accessed 12 Mar 2018

24. K. Hartung, Computer-on-modules for robotics & industry 4.0 automation. Combined Print Magazine for the European Embedded Market, pp. 6–8 (2017), http://files.iccmedia.com/magazines/basoct17/basoct17.pdf. Accessed 14 Mar 2018
25. P. Hu, S. Dhelim, H. Ning, T. Qiu, Survey on fog computing: architecture, key technologies, applications and open issues. J. Netw. Comput. Appl. **98**, 27–42 (2017). https://doi.org/10.1016/j.jnca.2017.09.002
26. J-M. Huet, Smart cities: the key to Africa's third revolution. BearingPoint (2016), https://www.bearingpoint.com/files/smart-cities-the-key-to-africas-third-revolution.pdf. Accessed 22 Feb 2018
27. R.L. Krutz, R.D. Vines, *Cloud security: A Comprehensive Guide To Secure Cloud Computing* (Wiley Publishing, Indianapolis, 2010)
28. M. Liyanage, C. Chang, S. N. Srirama, mePaaS: Mobile-embedded platform as a service for distributing fog computing to edge nodes. 2016 17th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), Guangzhou, China, 16–18, pp. 73–80 (Dec 2016). https://doi.org/10.1109/PDCAT.2016.030
29. P. Mach, Z. Becvar, Mobile edge computing: a survey on architecture and computation offloading. IEEE Commun. Surv. Tutorials **19**(3), 1628–1656 (2017). https://doi.org/10.1109/COMST.2017.2682318
30. Y. Mao, C. You, J. Zhang, K. Huang, K.B. Letaief, A survey on mobile edge computing: the communication perspective. IEEE Commun. Surv. Tutorials **19**(4), 2322–2358 (2017). https://doi.org/10.1109/COMST.2017.2745201
31. R. Minerva, A. Biru, D. Rotondi, Towards a definition of the internet of things (IoT). IEEE Internet Things, Revision 1 (2015), https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf. Accessed 14 Mar 2018
32. J. Mundy, 5G in the smart city (2018), https://5g.co.uk/guides/5g-smart-city/. Accessed 12 Mar 2018
33. Nebbiolo Technologies, Fog computing: keystone of industry 4.0 (2017), https://www.nebbiolo.tech/wp-content/uploads/Nebbiolo-Technologies-solutions-brief.pdf. Accessed 10 Mar 2018
34. N. Pokhriyal, W. Dong, V. Govindaraju, *Big Data for Improved Diagnosis of Poverty: A Case Study of Senegal* (The Brookings Institute, Washington, 2015), https://www.brookings.edu/blog/africa-in-focus/2015/06/02/big-data-for-improved-diagnosis-of-poverty-a-case-study-of-senegal/. Accessed 4 Apr 2018
35. P.C. Neves, J. Bernardino, Big data in the cloud: a survey. Open J. Big Data (OJBD) **1**(2), 1–18 (2015)
36. Nexcom, Industry 4.0 smart services dawn with predictive maintenance (2015), http://www.nexcom.com/news/Detail/industry-4-0-smart-services-dawn-with-predictive-maintenance?preview=1. Accessed 14 Mar 2018
37. Z. Pang, L. Sun, Z. Wang, E. Tian, S.Yang A survey of cloudlet based mobile computing. 2015 International Conference on Cloud Computing and Big Data (CCBD), Shanghai, pp. 268–275 (2015). https://doi.org/10.1109/CCBD.2015.54
38. C. Perera, Y. Qin, J. Estrella, S. Reiff-Marganiec, A. Vasilakos, Fog computing for sustainable smart cities: a survey. ACM Comput. Surveys (CSUR) **50**(3), 43 (2017). https://doi.org/10.1145/3057266
39. J.S. Preden, K. Tammemae, A. Jantsch, M. Leier, A. Riid, E. Calis, The benefits of self-awareness and attention in fog and mist computing. IEEE Comput. Soc. Comput. **48**(7), 37–45 (2015)
40. J. Preden, Evolution of mist computing from fog and cloud computing THINNECT (2014), http://www.thinnect.com/static/2016/08/cloud-fog-mist-computing-062216.pdf. Accessed 15 Mar 2018
41. I. Protopop, A. Shanoyah, Big data and smallholder farmers: big data applications in the agri-food supply chain in developing countries. Int. Food Agribusiness Manag. Rev. **19**(A), 173–190 (2016)

42. J.C. Ramirez, Fog, mist and hardware in factory automation (2017), https://www.smartindustry.com/blog/smart-industry-connect/fog-mist-and-hardware-in-factory-automation/. Accessed 14 Mar 2018
43. J.C. Ramirez, Hardware for industrial IoT fog and mist computing. Combined print magazine for the European Embedded Market, pp. 24–25 (2017), http://files.iccmedia.com/magazines/basoct17/basoct17.pdf. Accessed 14 Mar 2018
44. P.P. Ray, An introduction to dew computing: definition, concept and implications. IEEE Access **6**, 723–737 (2018). https://doi.org/10.1109/ACCESS.2017.2775042
45. B.P. Rimal, E. Choi, I. Lumb, A taxonomy and survey of cloud computing. 2009 Fifth International joint conference on INC, IMS and IDC, Seoul, pp. 44–51 (2009). https://doi.org/10.1109/NCM.2009.218
46. A. Rindos, Y. Wang, Dew computing: The complementary piece of cloud computing. 2016 IEEE International conferences on big data and cloud computing (BDCloud), Social computing and networking (SocialCom), Sustainable computing and communications (SustainCom) (BDCloud-SocialCom-SustainCom), Atlanta, GA, pp. 15–20 (2016). https://doi.org/10.1109/BDCloud-SocialCom-SustainCom.2016.14
47. R. Saracco, Industry 4.0 leveraging on 5G (2017), http://sites.ieee.org/futuredirections/2017/09/19/industry-4-0-leveraging-on-5g/. Accessed 12 Mar 2018
48. M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, The case for VM-based cloudlets in mobile computing. IEEE Pervasive Comput. **8**(4), 14–23 (2009). https://doi.org/10.1109/MPRV.2009.82
49. M. Satyanarayanan, Z. Chen, K. Ha, W. Hu, W. Richter and P. Pillai, Cloudlets: at the leading edge of mobile-cloud convergence. 6th International conference on mobile computing, applications and services, Austin, pp. 1–9 (2014). https://doi.org/10.4108/icst.mobicase.2014.257757
50. Siemens Discover MindSphere (2018), https://siemens.mindsphere.io. Accessed 14 Mar 2018
51. Siemens MindSphere, the internet of things (IoT) solution (2018), https://www.siemens.com/global/en/home/products/software/mindsphere.html. Accessed 8 Mar 2018
52. K. Skala, D. Davidovic, E. Afgan, I. Sovic, Z. Sojat, Scalable distributed computing hierarchy: cloud, fog and dew computing. Open J. Cloud Comput. **2**(1), 16–24 (2015)
53. T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, D. Sabella, On multi-access edge computing: a survey of the emerging 5G network edge cloud architecture and orchestration. IEEE Commun. Surv. Tutorials **19**(3), 1657–1681 (2017)
54. R. Taylor, B. Frank, Big data opening R&D opportunities in Africa (2017), http://www.scidev.net/index.cfm?originalUrl=/sub-saharan-africa/data/opinion/big-data-opening-opportunities-africa.html&. Accessed 4 Apr 2018
55. A. Varghese, D. Tandur Wireless requirements and challenges in industry 4.0. 2014 International conference on contemporary computing and informatics (IC3I), Mysore, India, pp. 634–638 (2014). https://doi.org/10.1109/IC3I.2014.7019732
56. B. Varghese, R. Buyya, Next generation cloud computing: new trends and research directions. Futur. Gener. Comput. Syst. **79**(3), 849–861 (2018). https://doi.org/10.1016/j.future.2017.09.020
57. A. Vork, J. Tamm LED Lights with smart controllers. SmartEnCity project, Estonia, Cityntel (2016), https://smartencity.eu/about/solutions/led-lights-with-smart-controllers-tartu/. Accessed 12 Mar 2018
58. Y. Wang, Cloud-dew architecture. Int. J. Cloud Comput. **4**(3), 199–210 (2015)
59. Y. Wang, Definition and categorization of dew computing. Open J. Cloud Comput. **3**(1), 1–7 (2016)
60. Y. Wang, Y. Pan, Cloud-dew architecture: realizing the potential of distributed database systems in unreliable networks. 21th International conference on parallel and distributed processing techniques and Applications (PDPTA'15), pp. 85–89 (2015)

Chapter 8
IoT and Edge Computing as a Tool for Bowel Activity Monitoring

Umit Deniz Ulusar, Erdinc Turk, Ahmet Sefa Oztas, Alp Erkan Savli, Guner Ogunc, and Murat Canpolat

8.1 Introduction

The Internet of Things (IoT) is enabling several applications ranging from environment to healthcare and medical fields [1, 2]. The main idea is the ubiquitous presence around us with a variety of things or objects that are connected to the Internet. A key application for the IoT-based technologies is wireless medical sensor networks (WMSNs), which aim to observe patients' health status remotely without intervening his/her daily life and perform early diagnosis of possible diseases. For this, researchers have been developing tiny sensors that can observe heartbeat, blood pressure, temperature, and so on, apply signal processing algorithms, and then share it through wireless connectivity with medical professionals. WMSNs are also an indispensable tool for patients with certain disabilities and accessibility issues by making remote monitoring possible and minimizing the necessity of hospital visits. Also, they are beneficial for healthy and active people for monitoring their daily activities and well-being. A typical example of this is a fall detection device which can report the situation to an emergency responder.

Unfortunately, so far, the real-world application of IoT-based health monitoring for mobile users suffered from energy constraints imposed by the batteries.

U. D. Ulusar (✉) · E. Turk
Department of Computer Engineering, Akdeniz University, Antalya, Turkey
e-mail: umitulusar@akdeniz.edu.tr

A. S. Oztas · A. E. Savli
Department of Electrical and Electronics Engineering, Akdeniz University, Antalya, Turkey

G. Ogunc
Department of General Surgery, Akdeniz University, Antalya, Turkey

M. Canpolat
Department of Biophysics, Akdeniz University, Antalya, Turkey

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_8

Typically, energy consumption for IoT devices is maximized during wireless communication. Edge computing aims to process data produced by devices to be closer to its origin instead of sending it to data centers.

In this chapter, we present a gastrointestinal track motility monitor (GTMM), an IoT-driven eHealth device that is specifically designed for constant monitoring of hospitalized patients after major abdominal surgery whose physiological status requires close attention. Similarly, this type of solution employs sensors to collect physiological information and processors for real-time data processing and uses gateways to send the data or warnings to caregivers for further analysis.

8.2 Background

Auscultation is the typical mean of observing body sounds and is commonly used in clinical environments. This technology exists since the early 1800s, and despite all the improvements, the overall aim is to identify specific sounds within the body. With the advent of technology, IoT-driven bioacoustic sensors can be utilized for continuous monitoring. There are some studies in the literature that aim to use bioacoustic sensors for observing pulmonary, heart, and intestinal sounds. Tang et al. have developed an electronic stethoscope for observing heart sounds with embedded processor and used Bluetooth for data transmission [3]. Similarly, Rach et al. have developed a bioacoustic sensor prototype that can operate over long periods and performed sound analysis [4]. In another research, Patil et al., have developed an electronic stethoscope to observe heart rhythm [5].

Electronic stethoscopes are simple devices, and they are typically designed for observing body sounds for short periods. In order to make them suitable for long-term monitoring, a detection algorithm for the sound type of interest is required. Various researchers have studied automatic and semiautomatic detection of intestinal sounds. Some techniques such as fractal dimension [6], wavelet decomposition [7], and neural networks [8] are used for short-term intestinal sounds analysis. Other studies by Dimoulas et al. [9], Kim et al. [10], and Emoto et al. [11] focused on more subtle change detections observed over longer durations.

For long-term monitoring of mobile users, another problem is the varying noise that exists at every level and at every location of our daily lives. Two types of noise require extra attention: environmental noise and body sounds. Some studies such as Hadjileontiadis et al. [12, 13] and Dimoulas et al. [14] focused on signal denoising and wavelet-based noise reduction. Jatupaiboon et al. [15] developed an electronic stethoscope prototype with adaptive noise cancellation. The proposed prototype acquires cardiac sounds and noises with two different microphones integrated in the stethoscope. LMS algorithm for adaptive filtering is applied. Jiao et al. [16] proposed a gradient adaptive step size LMS algorithm that uses two adaptive filters to estimate gradients more accurately, for biomedical applications. Some researches such as Ulusar [17] utilized classification techniques such as naive Bayesian and minimum statistics and spectral subtraction for noise attenuation.

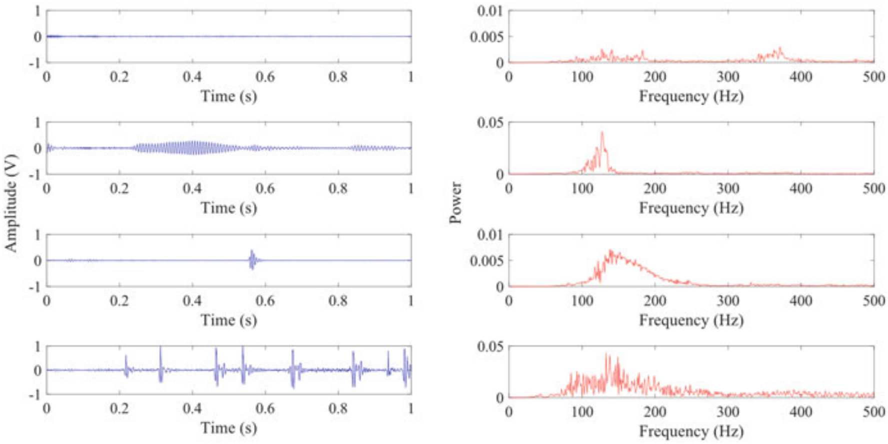


Fig. 8.1 Bowel sounds (quiet period, noise, single burst, multiple burst) and their corresponding power spectrum

8.2.1 Intestinal Sounds

The intestinal sounds (IS) refer to sounds produced within the small and large intestines during digestion. They are typically characterized as a nonstationary burst like signal and described as gurgling, rumbling, growling, and high pitched. These sounds are occasionally contaminated by other artifacts such as heartbeat, movement, breathing, and environmental noise [18]. Sound level varies and correlates with the intensity and the location of the bowel movement. Bowel sound (BS) is typically classified into two types: single burst (SB) and multiple bursts (MS) [18, 19]. Duration of single BS varies between 0.02 and 0.1 s and creates a spiky sound. A detailed work on abdominal sound classification can be found in [20]. Figure 8.1 depicts some of the observed intestinal sounds.

8.3 Methodology

8.3.1 Bioacoustic Sensor

Two electret microphones (POM-5246P-R) are used for bowel activity monitoring. The low-amplitude signals of the microphones were amplified with operational amplifier (LM358) and were digitized using a 10-bit ADC at a sampling frequency of 2048 Hz. The bias voltage of the operational amplifier circuit was set to be 1.55 V. The system's operating voltage was 3.3 V. Figure 8.2 shows the sensing device design.

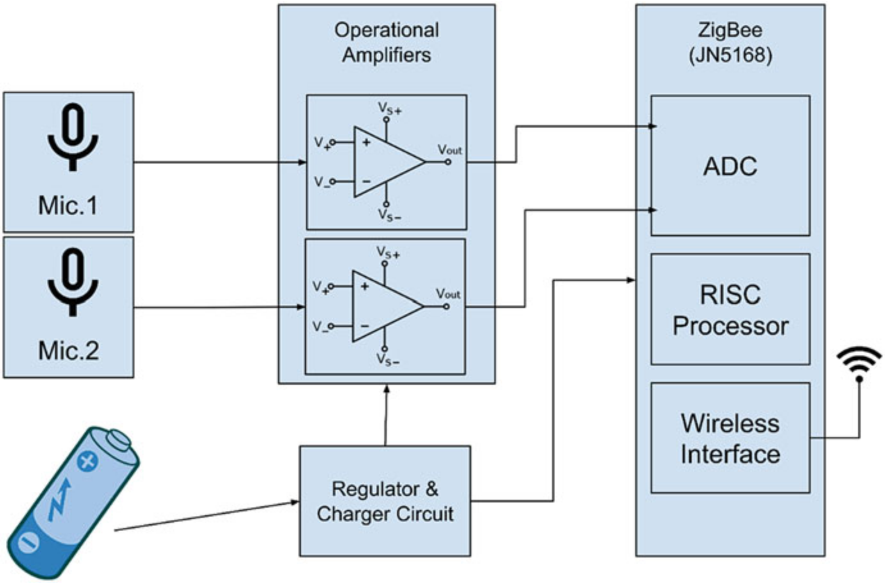


Fig. 8.2 IoT-driven sensor device

There are several methods of wireless communication such as Bluetooth, Wi-Fi, ZigBee, and GSM. Bluetooth is a major technology for wearable products and more suitable for transferring small chunks of data in short distances. Wi-Fi is widely used, and it provides serious data throughput in the ranges of Gbps. Typically, this type of communication is not suitable for IoT applications because of high power consumption. GSM uses various frequency ranges, and communication distance can go up to a couple of hundred kilometers, but due to its energy consumption, it is only suitable for low bandwidth data transfers. ZigBee is based on the IEEE802.15.4 protocol, operating at 2.4 GHz and suitable for short to mid-range applications. It offers high security, robustness, and high scalability. Due to energy consumption and data throughput rate, ZigBee module JN5168 (NXP, Sheffield, UK) is used for ADC and wireless communication. It has a RISC 32-bit processor and 4 kB EEPROM, 32 kB RAM, and 256 kB flash. This module consumes less than 0.12 μA during deep sleep state and less than 25 mA during communication state.

The open-circuit voltage changes with charging state of the battery. This variability is an undesirable condition for operational amplifier output, and a voltage regulator is used to generate 3.3 V fixed supply voltage. Additional capacitors are used to eliminate noise. Figure 8.3 shows the circuit design of the low noise voltage regulator design.

Figure 8.4 shows the circuit diagram of the battery charging module. A micro-USB jack is used for power connection and battery charging, and full situations are indicated with red and green LEDs.

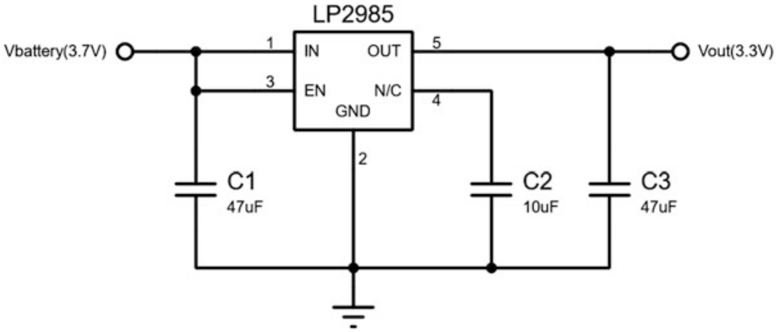


Fig. 8.3 Regulator circuit scheme

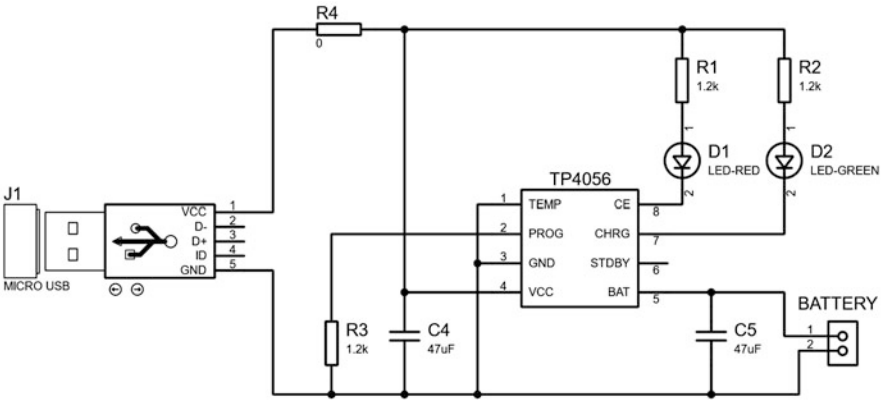


Fig. 8.4 Battery charger circuit

Figure 8.5 shows the one circuit diagram of operational amplification unit applied on both channels.

Figure 8.6 shows the custom box which was designed to hold the microphones, electronic parts, and the battery. In order to amplify body sounds, the bottom part has a conic shape similar to a stethoscope chest piece. Microphone 1 is located inside the stethoscope's conic metal part in order to obtain intestinal sounds. Microphone 2 is set to the upper side of the bioacoustic sensor to obtain ambient noise. Covering box, made by a 3D printer, is utilized to hold the electronic circuit and battery.

8.3.2 Signal Processing Steps

The sound intensity drops with distance, and microphone orientation plays an important role on which sounds will be acquired. By utilizing two microphones, we aimed to observe environmental sounds with one microphone and the body

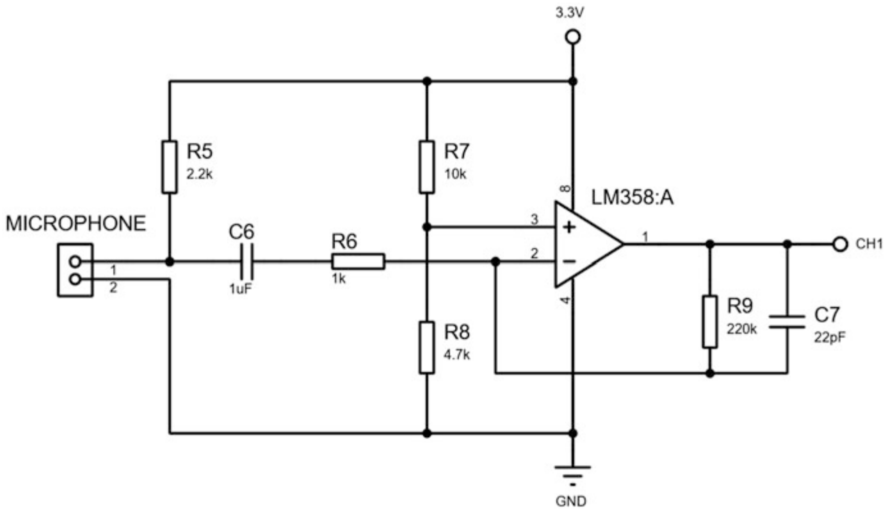


Fig. 8.5 Operational amplifier circuit

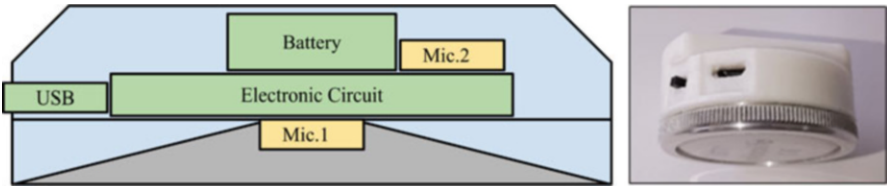


Fig. 8.6 Bioacoustic sensor mechanical design

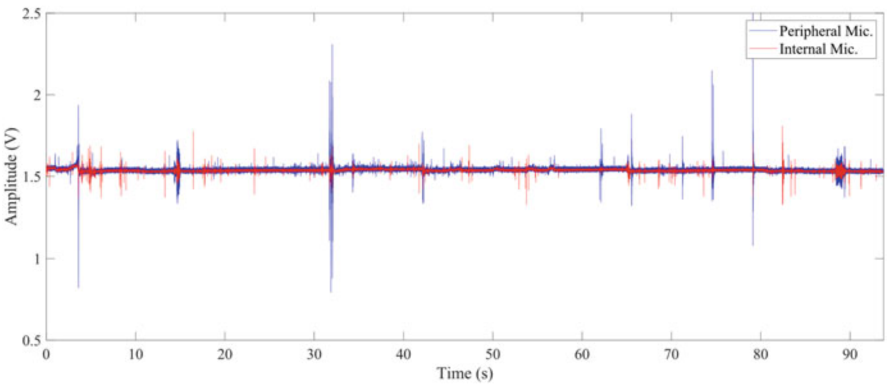


Fig. 8.7 90-s recordings from the internal and peripheral microphone

sounds with the other. The internal microphone (Mic.1) observes bowel sounds and ambient noise. Peripheral microphone (Mic.2) observes only ambient noise. Figure 8.7 shows 90-s-long sample recording. The blue line indicates the signals observed

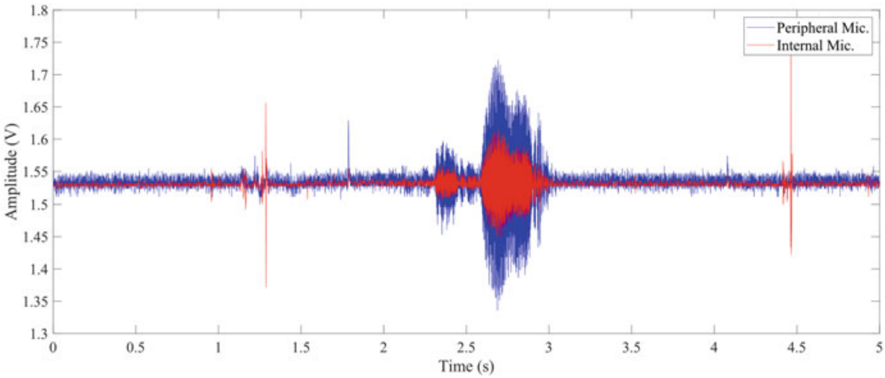


Fig. 8.8 5-s-long sample from the entire signal

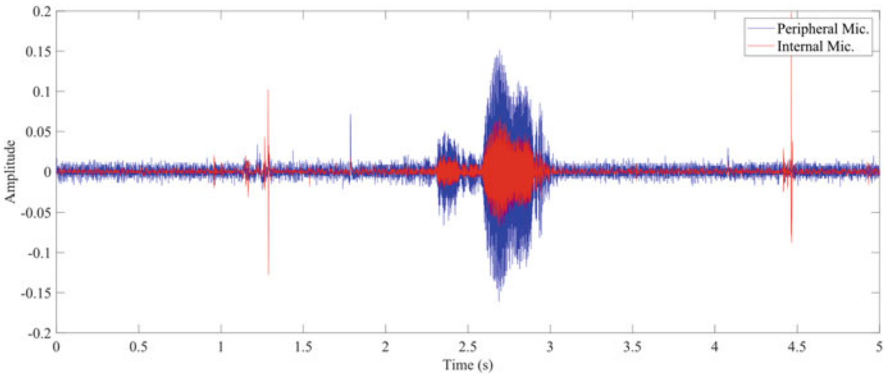


Fig. 8.9 5-s-long signal after high-pass filter and normalization

by the peripheral microphone, and the red line shows the internal microphone. The effect of the environmental noise is visible in both recordings, but the intensities vary.

Figure 8.8 shows a closer look of the observed noise by the microphones. Single bursts with high intensities are observed at 1.2 and 4.5 s.

The ZigBee module has a 500 MHz RISC-based processing unit and 292 kB internal memory. Most of the signal processing is performed on the edge device to minimize transferred data. Only signal chunks detected as activity (intestinal or not) are transferred for further processing. Data is processed as 10-s chunks buffered in the edge device's memory. Initially, 10-Hz high-pass filter is applied to remove DC offset, and signals are normalized (Fig. 8.9).

Hilbert transform-based envelope is used on both signals, and activity regions are detected by using an empirically chosen threshold value (0.02). Very short threshold crossings (<0.05 s) are discarded. Figure 8.10 shows the application of the Hilbert envelope on both signals.

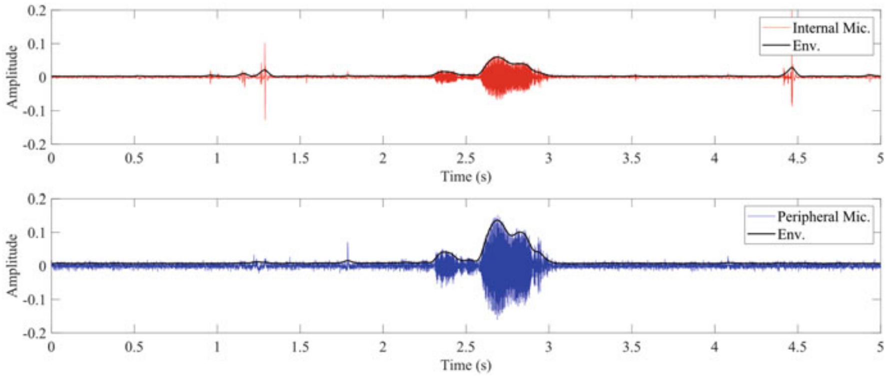


Fig. 8.10 Hilbert envelope and application on signals from the internal and peripheral microphone. The black solid line shows the temporal envelope obtained using Hilbert transform

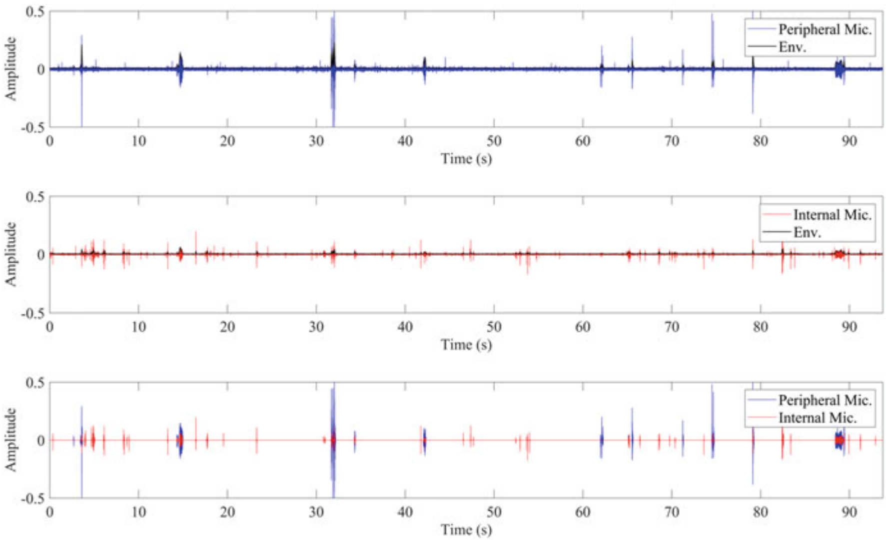


Fig. 8.11 Entire signal with envelopes

The start and the end time of activities and the signal above thresholds is transferred to the edge computing unit with more processing power and memory for further processing. In our case, this was a personal computer. Figure 8.11 shows the signals obtained from both microphones and the data transferred to the edge computing unit. There were 19 signal chunks for the peripheral microphone and 53 signal chunks for the signal obtained from internal microphone.

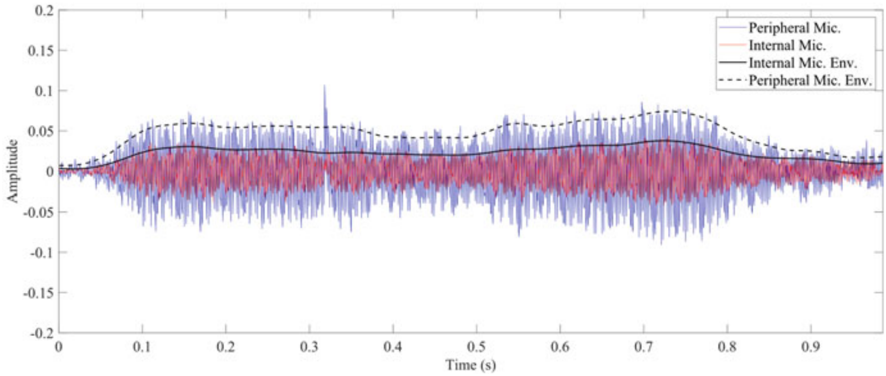


Fig. 8.12 Noise chunk observed by both microphones

8.3.3 Intestinal Activity Detection

Intestinal activities are detected by using rule-based and spectrum subtraction-based detection algorithm. In statistics, the outlier values are typically detected by using the mean and the standard deviation of the data. Similarly, we used the standard deviation of the signal (updated when a signal chunk is received) of the peripheral microphone. If an activity is observed by both microphones and the intensity of the signal obtained from the peripheral microphone is more than 3 times the standard deviation of the noise signal, we assumed that the activity observed in internal microphone is due to the external noise and called that activity as non-BS. Figure 8.12 shows an example of a signal chunk where environmental noise was observed by both microphones.

The amplitude and the duration of the intestinal sounds vary between each patient. The signal is nonstationary, and in order to identify BSs, we utilized spectral features. Figure 8.13 shows the average spectrogram of the signals. Hamming window is applied on both signals to extract power spectrograms. Window length is set to 0.12 s with 50% overlap which is greater than the possible longest single burst intestinal activity (0.1 s). The effect of noise on both spectrums around 400 Hz is clearly visible.

Around 97.5% of the bowel sounds' signal energy is located between 100 and 500 Hz [19]. For each chunk, spectral centroid, spectral bandwidth, and subband normalized energy were estimated from the frequency spectrum between 100 and 500 Hz. We used naive Bayesian approach with Kernel density estimation [21, 22] for the classification of chunks. For NB-based classification, the features are assumed statistically independent. With observed feature values $x_1, \ldots, x_n$, the posterior probability of each class c is estimated as follows:

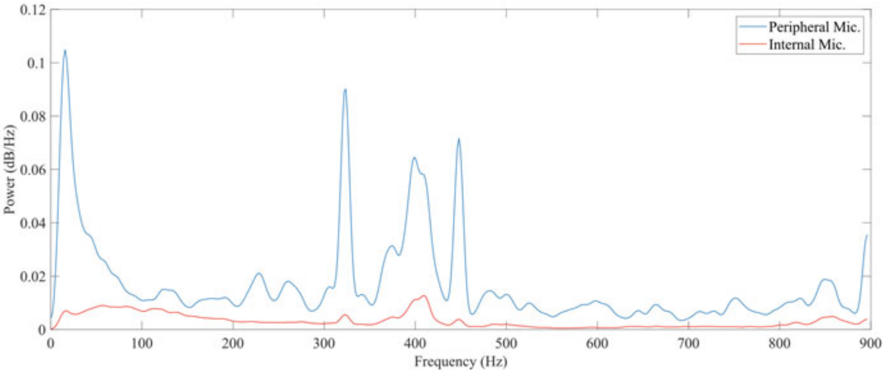


Fig. 8.13 Power spectrogram of signals obtained by microphones

$$p(x, c) = p(c) \prod_{i=1}^{D} p(x_i|c)$$

The class with the maximum posterior probability is assigned to the frame. During training, an expert marks the chunks into two groups, BS and non-BS. The algorithm learns the probability distribution of each class $p(c)$ and the class dependent feature distribution $p(x_i|c)$.

Finally, signals marked as BS are sent for further analysis. The intensity, frequency, and the duration of the signals are used for long-term bowel activity monitoring and abnormality detection.

8.4 Conclusion

IoT-related healthcare systems are low-powered networked devices that can collect and share biophysical data over time and provide intelligent decision for preventing serious illnesses. With the help of clinical decision support software, they have the power to transform healthcare and provide more precise and timely decisions.

This work introduces an edge computing and IoT-driven monitoring system for bowel sounds. Signal chunks of intestinal activity is marked and transferred to cloud-based analysis system. The advantages of edge computing are utilized by avoiding the need for energy expensive raw data communication.

Acknowledgment This work is supported by Akdeniz University Scientific Research Projects Coordination Unit (FYL-2015-1043).

References

1. F.M. Al-Turjman, Towards smart ehealth in the ultra large-scale Internet of Things era, in *2016 23rd Iranian Conference on Biomedical Engineering and 2016 1st International Iranian Conference on Biomedical Engineering (ICBME)*, (2016), pp. 102–105
2. U.D. Ulusar, F. Al-Turjman, G. Celik, An overview of Internet of things and wireless communications, in *2017 International Conference on Computer Science and Engineering (UBMK)*, (2017), pp. 506–509
3. Y. Tang, G. Cao, H. Li, K. Zhu, The design of electronic heart sound stethoscope based on bluetooth, in *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, (2010), pp. 1–4
4. M.M. Rach, H.M. Gomis, O.L. Granado, M.P. Malumbres, A.M. Campoy, J.J.S. Martín, On the design of a bioacoustic sensor for the early detection of the red palm weevil. Sensors **13**(2), 1706–1729 (2013)
5. D.D.K. Patil, R.K. Shastri, Design of wireless electronic stethoscope based on zigbee. Int. J. Distrib. Parallel Syst. **3**(1), 351–359 (2012)
6. L.J. Hadjileontiadis, I.T. Rekanos, Detection of explosive lung and bowel sounds by means of fractal dimension. IEEE Signal Process. Lett. **10**(10), 311–314 (2003)
7. L.J. Hadjileontiadis, C.N. Liatsos, C.C. Mavrogiannis, T.A. Rokkas, S.M. Panas, Enhancement of bowel sounds by wavelet-based filtering. I.E.E.E. Trans. Biomed. Eng. **47**(7), 876–886 (2000)
8. K.-S. Kim, J.-H. Seo, C.-G. Song, Non-invasive algorithm for bowel motility estimation using a back-propagation neural network model of bowel sounds. Biomed. Eng. Online **10**, 69 (2011)
9. C. Dimoulas, G. Kalliris, G. Papanikolaou, A. Kalampakas, Long-term signal detection, segmentation and summarization using wavelets and fractal dimension: A bioacoustics application in gastrointestinal-motility monitoring. Comput. Biol. Med. **37**(4), 438–462 (2007)
10. K.S. Kim, J.H. Seo, S.H. Ryu, M.H. Kim, C.G. Song, Estimation algorithm of the bowel motility based on regression analysis of the jitter and shimmer of bowel sounds. Comput. Methods Prog. Biomed. **104**(3), 426–434 (2011)
11. T. Emoto et al., ARMA-based spectral bandwidth for evaluation of bowel motility by the analysis of bowel sounds. *Physiol. Meas.***34**(8), 925 (2013)
12. L.J. Hadjileontiadis, Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding-part II: Application results. I.E.E.E. Trans. Biomed. Eng. **52**(6), 1050–1064 (2005)
13. L.J. Hadjileontiadis, Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding-part I: Methodology. I.E.E.E. Trans. Biomed. Eng. **52**(6), 1143–1148 (2005)
14. C. Dimoulas, G. Kalliris, G. Papanikolaou, A. Kalampakas, Novel wavelet domain Wiener filtering de-noising techniques: Application to bowel sounds captured by means of abdominal surface vibrations. Biomed. Signal Process. Control **1**(3), 177–218 (2006)
15. N. Jatupaiboon, S. Pan-ngum, P. Israsena, Electronic stethoscope prototype with adaptive noise cancellation, in *2010 Eighth International Conference on ICT and Knowledge Engineering*, (2010), pp. 32–36
16. Y. Jiao, R.Y.P. Cheung, W.W.Y. Chow, and M.P.C. Mok, A novel gradient adaptive step size LMS algorithm with dual adaptive filters, in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (2013), pp. 4803–4806
17. U.D. Ulusar, Recovery of gastrointestinal tract motility detection using Naive Bayesian and minimum statistics. Comput. Biol. Med. **51**, 223–228 (2014)
18. C. Dimoulas, G. Kalliris, G. Papanikolaou, V. Petridis, A. Kalampakas, Bowel-sound pattern analysis using wavelets and neural networks with application to long-term, unsupervised, gastrointestinal motility monitoring. Expert Syst. Appl. **34**(1), 26–41 (2008)

19. R. Ranta, V. Louis-Dorr, C. Heinrich, D. Wolf, F. Guillemin, Digestive activity evaluation by multichannel abdominal sounds analysis. I.E.E.E. Trans. Biomed. Eng. **57**(6), 1507–1519 (2010)
20. C.A. Dimoulas, G.V. Papanikolaou, V. Petridis, Pattern classification and audiovisual content management techniques using hybrid expert systems: A video-assisted bioacoustics application in abdominal sounds pattern analysis. Expert Syst. Appl. **38**(10), 13082–13093 (2011)
21. N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F Radar Signal Process.***140**(2), 107 (1993)
22. E. Frank, L. Trigg, G. Holmes, I.H. Witten, Technical note: Naive Bayes for regression. Mach. Learn. **41**(1), 5–25 (2000)

Chapter 9
Application of Cloud Computing and Internet of Things to Improve Supply Chain Processes

S. Kanimozhi Suguna and Suresh Nanda Kumar

9.1 Introduction to IoT and Cloud Computing

In today's hypercompetitive environments in which competition is between supply chains and not between firms, individual firms need to effectively manage increasingly global and extending supply chain activities. The supply chain has become the central organizing unit in today's global industries [48]. As the Internet of Things (IoT) is opening up new business opportunities, firms are trying to understand the impacts of the IoT revolution on their supply chain. IDC [31] predicts that the IoT base will grow at a 17.5% compound annual growth rate (CAGR), reaching 28.1 billion installed units by 2020, up from 9.1 billion in 2013, and the value of the market will reach \$7.1 trillion by 2020, up from \$1.9 trillion in 2013.

The IoT has a significant impact on several aspects of everyday life and behavior of potential users [5]. Assisted living, smart homes, e-health, and enhanced learning are a few examples of possible application scenarios in which the IoT will play a leading role in the near future [4]. The IoT will provide new opportunities for services and applications able to leverage the interconnection of physical and virtual realms [50], and firms in most industries will rapidly adopt IoT-enabled applications in order to stay competitive.

Deutsche Bahn, the German Railways and cargo carrier, installed a network-wide monitoring system to manage its entire rail network which comprises over 1 billion supply chain "nodes," collecting data on each segment of track, rail car, station, engine, and switch and monitoring the condition of all of these things in real time.

S. Kanimozhi Suguna
Department of CSE, School of Computing, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India

S. Nanda Kumar (✉)
CII Institute of Logistics, Chennai, Tamil Nadu, India

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_9

The collected data are fed into a control tower that aggregates them every 5 s to provide near-real-time information across the entire fleet. Deutsche Bahn has used these data to improve risk management practices such as real-time rerouting and optimization, considering all existing network traffic through nodes [45]. Whirlpool is another example of using the IoT for internal supply chain optimization in routing work and locating misplaced inventory. Instead of using bar codes or a similar solution, Whirlpool used radio-frequency identification (RFID) tags and readers across a manufacturing plant to give managers and operators real-time access to information for inbound logistics to the paint line [45].

According to the Gartner Press Release [21], the following technologies will be extensively used in the coming years to become digital businesses. These technologies are Bioacoustic Sensing; Digital Security; Smart Workspace; Connected Home; 3D Bioprinting Systems; Affective Computing; Speech-to-Speech Translation; Internet of Things; Cryptocurrencies; Wearable User Interfaces; Consumer 3D Printing; Machine-to-Machine Communication Services; Mobile Health Monitoring; Enterprise 3D Printing; 3D Scanners; Consumer Telematics. The IoT can be used by the various partners of the supply chain to monitor the supply chain execution process in real time and improve the efficiency and effectiveness of supply chain [53]. Recent developments in the IoT have made it possible to achieve high visibility in the supply chain. The location of arbitrary individual things can be determined at any point in time by all appropriate supply chain partners [22]. For example, the IoT benefits the food and agricultural product supply chain by improving the visualization and traceability of agriculture products and ensuring people's food safety [66]. Industrial deployment of the IoT provides development of an ideal platform for decentralized management of warehouses and collaborative warehouse order fulfillment with RFID, ambient intelligence, and multi-agent system [56].

Ravi and Arora [3] describe organizations as constantly engaged in actions. If an organization does not take actions, it cannot remain solvent, and if the organization consistently exercises poor judgment in the actions that it takes, it is destined to fail. Organizations do not have minds, and they do not make decisions, they are institutions that empower individual agents to make decisions on their behalf, and empower other individuals to carry out the actions associated with these decisions. These decision making agents can be individuals or groups. The complex nature of organizational decisions creates information overload for individual agents, causing them to engage in a number of information processing shortcuts, which threaten the quality of their decisions. The ability to access infrastructures, platforms, and software services based on what is needed and paying for only what is used put start-up enterprises and small and medium enterprises at an advantageous position vis-à-vis the competition from larger enterprises and firms. This also gives them an edge in the market and an almost an equal position with much larger firms and enterprises. This provides a level playing field and can put pressure on enterprises that have not yet realized the potential and impact that cloud computing can have when developing or implementing their strategies. There can also be pressure when the organizational culture is not open to emergent market opportunities and has a conservative approach to technology adoption that is best addressed by leveraging cloud solutions. When an enterprise does not consider the strategic implications of cloud computing on the existing and new market opportunities, it can result in areas of high pressure, and the following risks can be the outcome:

- New entrants can claim a part of the market share and thereby a segment of traditional market dominance.
- Strategies that do not address competitor capabilities.
- Less-than-expected benefits received from technology-dependent solutions.

Utility services and service supply chain enterprises have traditionally extracted value and advantage from the ownership of technology and systems. This value and advantage may be reduced in a market in which technology and systems can be easily and rapidly acquired, expanded, or contracted based on need and according to the financial resources available with the firm; development and management costs are shared among a large base of users; and cost is directly associated with resources used. In an environment in which computing is a technology utility, there is less concern for how the infrastructure, platform, or software services are developed. More focus is paid to organizational needs, expected benefits to be delivered, and the value that can be obtained from the computing utility. Here, the concept of agile enterprise comes into focus. The agile enterprise takes advantage and benefits from solutions that can be used as needed and discarded when they no longer provide value. Those enterprises are nimble and are highly responsive and sensitive to market opportunities and leverage supply infrastructures, platforms, and software services, creating a value supply chain that can be reconfigured as needs and market conditions change. These are highly adaptive and agile and are able to hedge risks by making use of technology resources such as infrastructure, platform, and software offered by external service providers. The view of computing as a utility and the delivery of cloud solutions as a supply chain of information system solutions put greater pressure on enterprises that contain a culture that is not accepting of utility solutions, a structure that does not facilitate cooperative planning and integration of utility solutions, and processes that cannot take advantage of computing solutions provided as a supply chain of utilities. Enterprises that fail to consider the impacts of cloud solutions seen as a computing utility brought together into a supply chain of IT capabilities may experience areas of pressure resulting in the following risk outcomes:

- Overinvestment of resources in planning and building internally developed information system solutions
- Less-than-optimal results when value-producing cloud utilities are missing from the total solution
- Duplication of effort when specialist services available through cloud providers are not integrated as part of system management
- Less-than-expected results when utility components are not integrated into and managed as an information system capability supply chain

9.2 Supply Chain Management

The supply chain itself is an arrangement of associations, individuals, exercises, data, and assets engaged with the arranging, moving, or capacity of an item or administration from provider to client (in reality more like a “web” than a “chain”). Inventory network exercises change regular assets, crude materials, and parts into a completed item that is conveyed to the end client.

Conversely, supply chain management, as characterized by the Council of Supply Chain Management Professionals (CSCMP), “incorporates the arranging and administration of all exercises engaged with sourcing, obtainment, transformation, and coordination administration. It likewise incorporates the vital segments of coordination and cooperation with channel accomplices, which can be providers, mediators, outsider specialist co-ops, and clients.”

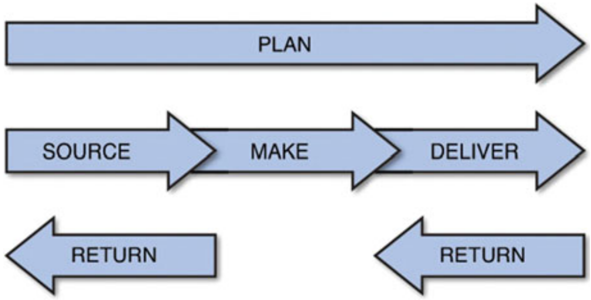
Fundamentally, supply chain management incorporates free market activity administration inside and crosswise over organizations and commonly “incorporates the majority of the coordination administration exercises noted above, and also fabricating operations, and it drives coordination of procedures and exercises with and crosswise over promoting, deals, item configuration, back and data innovation” [68].

Contingent upon one’s view, a portion of the capacities recorded here might be incorporated into the inventory network and coordination association:

- *Procurement*: The procurement of merchandise or administrations from an outside outer source
- *Demand determining*: Estimating the amount of an item or administration that clients will buy
- *Customer administration and request administration*: Tasks related to satisfying a request for products or administrations set by a client
- *Inventory*: Planning and administration
- *Transportation*: For contract and private
- *Warehousing:* Public and private
- *Materials dealing with and bundling*: Movement, insurance, stockpiling, and control of materials and items utilizing manual, semi-robotized, and mechanized hardware
- *Facility arrange*: Location choice in an association’s production network organized

Production network administration is likewise entwined with operations administration, which comprises exercises that make an incentive by changing sources of info (i.e., crude materials) into yields (i.e., merchandise and enterprises). The two exercises bolster the assembling procedure.

Fig. 9.1 SCOR model



9.2.1 SCOR Model

Another approach to see the production network is through the SCOR display, which was created by the Supply Chain Council (SCC) [69] to instruct, comprehend, and oversee supply chains. It is a model to both characterize and measure the execution of an association's production network.

The SCOR display is sorted out around the five noteworthy administration forms (see Fig. 9.1):

- *Plan*: Alignment of assets to request
- *Make*: Conversion or esteem included exercises inside a production network operation
- *Source*: Buying or procuring materials or administrations
- *Deliver*: All client collaboration, from getting a request to conclusive conveyance and establishment
- *Return*: All procedures that turn around material or administration streams from the client in reverse through the store network

This gives a wide definition to the inventory network, which features its significance to the association and how it makes measurements to gauge execution.

9.2.2 An Integrated Value-Added Supply Chain

The objective of the present store network is to accomplish combination through the joint effort to accomplish visibility downstream toward the client and upstream to providers. As it were, a large number of the present organizations have possessed the capacity to "substitute data for stock" to accomplish efficiencies. The times of having "islands of robotization," which may enhance association's inventory network at the cost of another person's (for instance, provider), are finished.

The ideas of collaboration and basic reasoning helped by innovation empower associations to work with different capacities inside and with different individuals from their store network, including clients, providers, and accomplices, to

Fig. 9.2 The value chain

accomplish new levels of productivity and to utilize their inventory network to accomplish an upper hand that spotlights on increasing the value of the client rather than simply being a cost focus inside the association.

The value chain model, began by Michael Porter, demonstrates the esteem making exercises of an association, which as should be obvious in Fig. 9.2 depend intensely on inventory network capacities.

In a valued chain, each of a company's inside exercises recorded after the figure increases the value of the last item or administration by changing contributions to yields.

- *Inbound coordinations*: Receiving, warehousing, and stock control of info materials
- *Operations*: Transforming contributions to the last item or administration to make esteem
- *Outbound coordinations*: Actions that get the last item to the client, including warehousing and request satisfaction
- *Showcasing and deals*: Activities identified with purchasers acquiring the item, including promotion, estimation, appropriation channel determination, and so forth
- *Administration*: Activities that keep up and enhance an item's esteem, including client bolster, repair, guaranteed benefit, and so forth

Bolster exercises distinguished by Porter can likewise increase the value of an association:

- *Obtainment*: Purchasing crude materials and different sources of info that are utilized as a part of significant worth making exercises
- *Innovation improvement*: Research and advancement, process computerization, and comparative exercises that help esteem chain exercises
- *Human asset administration*: Recruiting, preparing, improvement, and pay of representatives
- *Firm foundation*: Finance, legitimate, quality control, etc

Porter prescribed value chain analysis to research regions that speak to potential qualities that can be utilized to accomplish an upper hand. As should be obvious, the inventory network mixes it up of ways, so it ought to be a basic region of center [65].

Table 9.1 Supply chain leveraging effect

	Current	Supply chain improvement option	Sales increase option
Sales	$1,000,000	$1,000,000	$1,250,000
Cost of material	$650,000 (65%)	$600,000 (60%)	$812,500 (65%)
Production costs	$150,000 (15%)	$150,000 (15%)	$187,500 (15%)
Fixed costs	$100,000 (10%)	$100,000 (15%)	$100,000 (8%)
Profit	$100,000 (10%)	$150,000 (15%)	$150,000 (15%)

9.2.3 Levaraging the Supply Chain

Since inventory network costs speak to a critical segment of an organization's business, it is not hard to perceive any reason why there is such an emphasis on it. This outcome in a "utilizing" impact, as any dollar saved money on inventory network, contributes as the same to the primary concern as a considerably bigger and frequently unattainable increment in deals (will change in light of an individual organization's overall revenue).

Table 9.1 represents this through a case of a business that is assessing two key choices: (1) decrease its inventory network costs by roughly 6.5% through more compelling arrangements with a merchant, or (2) increment deals by 25% (which will mostly without a doubt likewise add to deals and promoting costs). The utilizing impact of the store network as the generally little cost diminishes contributes as much to the primary concern as the 25% deals increment (which is quite hard to achieve in any economy) is represented.

The production network cost decrease in this illustration has amazing outcomes; however, you need to remember that "you can't get blood from a stone." Through Lean strategy, a group-based type of constant change that spotlights on the recognizable proof and disposal of waste, it can make a "change in outlook" that can make process (and cost) upgrades that were already thought unimaginable.

9.2.4 Supply Chain Strategy for a Competitive Advantage

Production network and logistics capacities were seen fundamentally as cost focuses to be controlled. It is just in the previous 20 years or with the goal that it has turned out to be evident that it can be utilized for an upper hand also.

To achieve this, an association ought to set up aggressive needs that their production network must need to fulfill inner and outer clients. They should then connect the chosen aggressive needs to their store network and coordination forms.

Krajewski et al. [67] recommend breaking an association's aggressive needs into cost, quality, time, and adaptability capacity gatherings:

- *Cost methodology*: Focuses on conveying an item or administration to the client at the most reduced conceivable cost without yielding quality. Walmart has been the minimal effort pioneer in retail by working a productive store network.
- *Time system*: This technique can be as far as speed of conveyance, reaction time, or even item improvement time. Dell has been a prime case of a producer that has exceeded expectations at reaction time by amassing, testing, and delivering PCs in as meager as a couple of days. FedEx is known for quick, on-time conveyances of little bundles.
- *Quality procedure*: Consistent, brilliant merchandise or administrations require a solid, safe production network to convey on this guarantee. On the off chance that Sony had a substandard store network with high harm levels, it would not make any difference to the client that their gadgets are of the most elevated quality.
- *Adaptability technique*: Can come in different structures, for example, volume, assortment, and customization. A considerable lot of the present online business organizations, for example, Amazon, offer a lot of adaptability in a large number of these classifications.

Much of the time, an association may concentrate on more than one of these procedures, and notwithstanding when concentrating on one, it does not imply that they will offer disappointing execution on the others (just not "best in class" maybe).

9.3 Applications of IoT in Logistics and SCM

Figure 9.3 demonstrates the different IoT applications in the supply chain. Supply chain advancement is basic to practical upper hands for firms. Supply chain innovation has been defined as "a change (incremental or radical) within a supply chain network, supply chain technology, or supply chain process (or a combination of these) that can take place in a company function, within a company, in an industry or in a supply chain in order to enhance new value creation for the stakeholder [2]. Production network information will progressively be created by a system of sensors, RFID labels, meters, actuators, GPSs, and different gadgets and frameworks [11]. IBM [30] recommends the smarter supply chain of the future has the instrumentation, interconnectedness, and knowledge to anticipate, if not forestalled, disturbances before they happen. From generation floor and warehousing to the conveyance and store racking, the IoT empowers production network accomplices to make brilliant inventory network by giving continuous information and business knowledge for all accomplices in the inventory network. Firms will put resources into the IoT to build deceivability of materials stream, decrease the loss of materials, and lower circulation costs. This is delineated in Fig. 9.4.

As the IoT enters into center business forms and an expanding number of firms put resources into green store network, utilizing the IoT for keen production network will turn into a subject of awesome enthusiasm for the chiefs [40]. The

IoT for keen store network will produce an incentive in stock administration, preventive support, and transportation and permit multipath interchanges with the accomplices and clients, rich information examination, and opportune reactions to sudden occasions. Notwithstanding, the estimation of the IoT can be acknowledged just when directors completely comprehend what information the IoT gadgets gather, how the information ought to be handled, and when they should settle on the choices on store network issues. It is likewise essential to break down specialized and administrative security dangers of the IoT [35].

General engineering will have a noteworthy bearing on the field itself and should be researched [24]. For effective IoT-based production network administration, fundamental IoT engineering should be built up and refreshed persistently to commission and decommission different IoT resources. Reaidy et al. [56] propose an IOT foundation for synergistic distribution center request satisfaction in light of RFID, surrounding insight, and multi-operator framework. It comprises a physical gadgets layer, a middleware surrounding the stage, a multi-operator framework, and an undertaking asset arranging. It incorporates a base up approach with choice help systems, for example, self-association and transaction conventions.

For brilliant store network, the essential engineering incorporates four layers [40]: (1) question layer, (2) correspondence layer, (3) application layer, and (4)

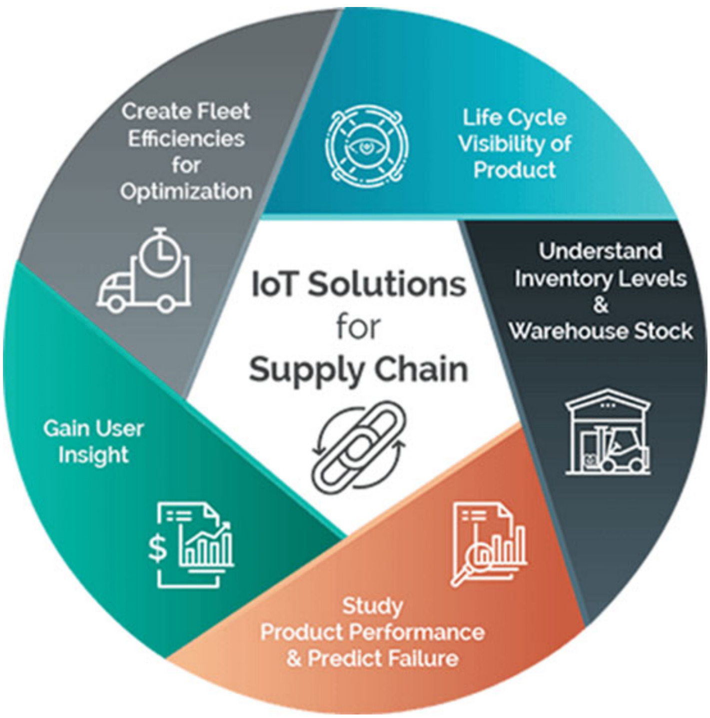


Fig. 9.3 The various IoT applications in the supply chain

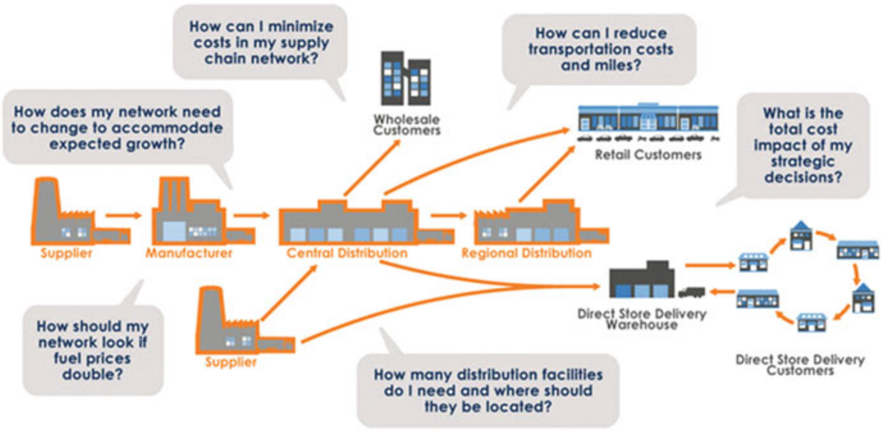


Fig. 9.4 Cost reduction and optimal distribution network design using IoT

information benefit layer. Each layer has center segments and basic functionalities depicted underneath: object layer comprises physical protests, for example, gadgets, machines, sensors, RFID labels, and peruser. The question layer is in charge of detecting nature, recognizing and following articles, and gathering information. The physical articles persuade scaled down to be more vitality productive, area autonomous, and financially savvy [59]. RFID labels and sensors are frequently installed in the machines and gadgets.

Application layer comprises an arrangement of issue particularly programming devices that associate with clients, take care of issues, process information, and offer arrangements with different applications. The application layer is in charge of displaying information and data to the clients in an easy to use arrange. While gadget-to-gadget applications do not really require information perception, more human-focused IoT applications give representation to exhibit data to end clients in a natural and straightforward route and to permit collaboration with the earth. It is critical for IoT applications to be worked with knowledge, so the gadgets helps in working with it. A standout among the essential results of the IoT is a gigantic measure of information created from gadgets associated with the Internet [24]. At the demand of the clients, a continuous examination of store network information is led for basic leadership. The cloud information benefit framework ought to have the capacity to powerfully organize demands and arrangement assets to such an extent that basic solicitations are served continuously [24]. Information mining innovations are utilized to find learning covered up in the ocean of information [10]. Not at all like the other three layers which are regularly possessed by an individual firm, the information benefit layer is frequently claimed and overseen by open cloud specialist organizations.

The IoT encourages the improvement of heaps of industry-situated and client-particular IoT applications [40]. Although gadgets and systems give the physical availability, IoT applications empower the gadget-to-gadget and human-to-gadget

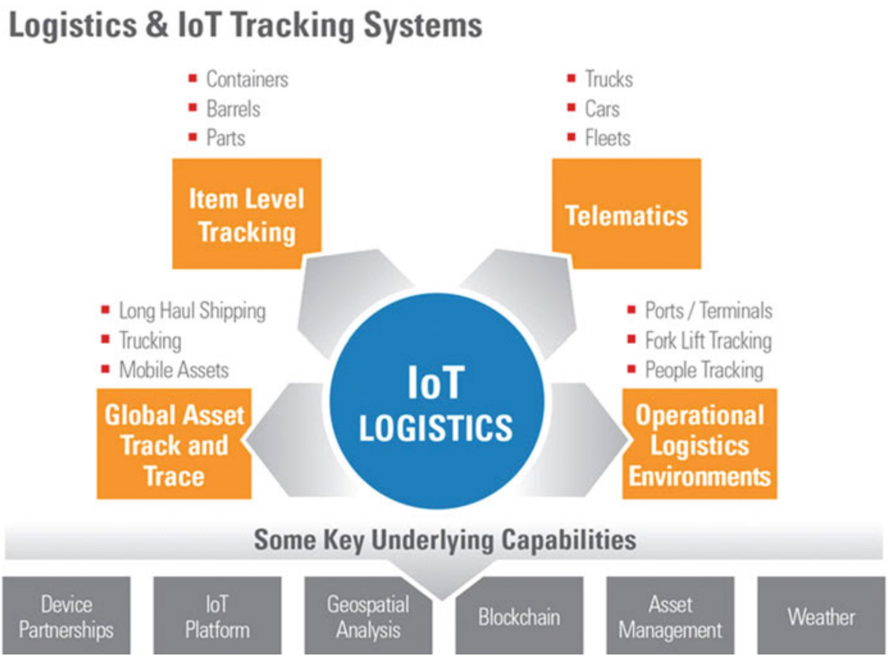


Fig. 9.5 Capabilities in IoT logistics

communications in a solid and vigorous way. IoT applications on gadgets need to guarantee that information/messages have been gotten and followed up on appropriately in an auspicious way. For instance, transportation and coordination applications screen the status of the shipped merchandise, for example, organic products, new cut create, meat, and dairy items. Amid the transportation, the preservation status (e.g., temperature, moistness, and stun) is checked always, and proper moves are made consequently to dodge waste when the status is out of range. For instance, FedEx utilizes SenseAware to watch the temperature, area, and other imperative indications of a bundle, including when it is opened and whether it was messed with the route. Figure 9.5 underneath indicates how IoT is utilized as a part of coordinations in thing level following, following vehicles utilizing telematics, following material taking care of gear like forklifts, and other whole deal following.

As the innovation advances, correspondence and handling capacities of the IoT are winding up increasingly open and flexible, and progresses in interconnectivity will drive the want to make utilization of the IoT's possibilities [59]. Rising innovation refers to innovation of a work in progress or at the phase of the verification of idea (POC), which has a high danger of specialized disappointment yet, however, has possibly noteworthy market esteem. For instance, checking and controlling innovation is moving to inconspicuous modest and vitality productive remote sensor innovation which enables applications to be sent all the more to adaptable areas autonomously. Portable crowd sensing is a promising advancement where people with detecting and processing gadgets on the whole offer information and concentrate data to gauge and guide marvels of regular intrigue [20]. The sensor arrangement is advancing toward a setting mindful of self-sorting out, self-repairing, and self-sufficient system. Innovation for data sharing and coordinated effort is moving to client arranged, conveyed insight framework, things-to-things joint effort situations, and things-to-people cooperation. For more secure data sharing and coordinated effort between gadgets, the rising innovation will reinforce interoperability, network, protection, and knowledge highlights of the IoT. The fate of enormous information and information investigation incorporates setting mindful information handling, subjective preparing, and enhancement [41].

Supply chain IoT gadgets create a huge measure of information that should be gathered and prepared progressively to give data on status, area, usefulness, and condition of items and administrations. Since a lot of sensor information continuously will build workloads of gadgets at an exponential rate, it is not practical to process all the gushing information accessible to those gadgets. Setting mindful information preparing empowers sensors and gadgets to utilize setting particular data, for example, area, temperature, and the accessibility of a specific gadget to choose what kinds of information to gather and decipher to give pertinent data to basic leadership. For instance, setting mindful information handling can convey significant data to a truck driver by knowing the truck's flow area (e.g., on an interstate, a congested inward city zone, or at a dangerous assembling plant). By incorporating the human perception process into IoT applications, an intellectual information handling application figures out how to detect things, foresee future occasions, and frame conclusions in view of confirmation under various conditions. For instance, intellectual information preparing utilizes computerized reasoning procedures to comprehend the warehousing condition, forms ecological information, for example, chemicals and contamination for a director, and uses an input from the administrator to learn further. The enhancement of information preparing is basic to an opportune procedure of the persistent stream of gigantic measures of information. Mechanical advances in streamlined information handling help settle on opportune choices in huge information applications in the keen store network. Figure 9.6 underneath demonstrates the different IoT organization structures amid the distinctive eras, past, present, and future.

IoT innovation stage is a mechanical establishment and improvement methodology with which each organization builds up their IoT applications for new esteem age and the upper hand. These stages are produced for the most part by IT firms and utilized by non-IT firms. IT firms set IoT models and build up an assortment of foundational advances for modern machines, ecological sensors, home apparatuses, and medicinal services gear. The decision of the IoT innovation stages by non-IT firms ought to be made regarding the general business and innovation methodology. As of now, there are four generally embraced innovation stages: exclusive IoT innovation, open development IoT innovation, industry union IoT innovation, and IoT as a Service (TaaS).

Exclusive IoT innovation has been created by many firms to procure an early-mover advantage, to end up plainly an innovation pioneer, and to be marked as a

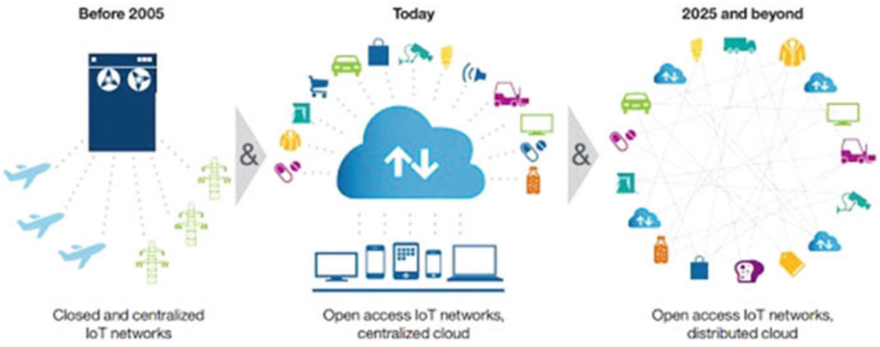


Fig. 9.6 IoT network architecture in different periods

trailblazer in their industry. For instance, Apple has its own licensed exclusive IoT innovation for in-home gadget correspondence, HomeKit. Exclusive IoT innovation is broadly produced for checking and control applications.

Open development IoT innovation is frequently utilized by a firm whose advances are constrained in making esteem included IoT applications. Open innovation has been viewed as applicable basically to “high-innovation” enterprises and has gotten more prominent consideration in light of the verbal confrontation about globalization and the potential for the R&D capacity to be outsourced [14]. Open-source IoT stages are favored and are more alluring contrasted and the exclusive IoT innovation stages because of the accompanying reasons. To begin with, the utilization of the open source is relied upon to empower the speedier combination of new IoT arrangements over the application spaces [49]. Second, the utilization of the open source has been accounted for to accelerate the appropriation of a product innovation in a base up form. At last, when seen from the social surplus point of view, the industry in view of the open-source stages has been found to give bigger aggregate welfare to the business structures [16].

Open advancement is known to quicken development process and make cooperative energy between endeavors of various innovation abilities. The main organization starts the open advancement and makes an open development coordination with other littler organizations which contribute their mechanical know-how to a particular issue. For instance, the Kimberly-Clark Digital Innovation Lab (D’Lab) conveyed open-source development to the buyer bundle merchandise class with the second yearly KChallenge start-up rivalry at the 2015 Consumer Electronics Show (CES) in Las Vegas [36]. The triumphant new businesses will get the chance to pilot a task with one of Kimberly-Clark’s worldwide brands, for example, Depends, Scott, and Kotex, to discover better approaches to interface with their customers. Kimberly-Clark fills in as an orchestrator and a facilitator of advancement, and taking an interest in new company assumes a basic part in propelling development for different brands of Kimberly-Clark.

Industry partnership IoT innovation has been made by driving IT firms to build up industry models and multi-gadget interoperability and to give clients esteem included administrations, for example, controlling various home systems administration gadgets. Driving IT firms make a consortium, joint wander, or organization. Google, through its in-home gadget unit Nest Labs, has propelled an industry organization working together with a few different firms to set a standard called thread for Internet-associated home gadget interchanges. Members of the business cooperation called the Thread Group incorporate Google, Samsung Electronics, and ARM Holdings. AllSeen Alliance, another gathering of more than 100 organizations including Microsoft and Qualcomm, was built up for the advancement and development of an interoperable associate network and interchange structure. The utilization of the business organization together with an open advancement is a decent improvement system for data sharing and coordinated effort applications.

Things as a Service (TaaS) is given as administration tasks like people in general cloud administrations. The rise of TaaS lies in the adaptation of unnecessary IoT limit, economies of scale and extent of shared IoT resources, pervasive availability by means of the Internet and remote systems, commoditization of IoT gadgets and applications, and industry institutionalization and interconnectivity. While TaaS imparts normal attributes to open cloud administrations, novel offers may originate from consolidating IoT structural parts, for example, physical items, arrangement, information stockpiling, and applications, and giving altered administrations to firms on a compensation as-you-go premise. For enormous information and information investigation, the industry organization together and TaaS are appropriate innovation stages.

9.4 Capabilities of IoT-Based Supply Chain Application

Based on the technology trends, this section discusses three capabilities of IoT-based supply chain applications: (1) monitoring and control, (2) information sharing and collaboration, and (3) big data and IoT data analytics.

9.4.1 Monitoring and Control

As per the Digital Agenda for Europe [17], observing and controlling allude to the control of any framework, gadget, or system through computerized techniques oversaw by a control unit with or without the capacity to show data. IoT-empowered checking and control frameworks are picking up fame in the natural and assembling areas, where organizations have worked fundamental applications, for example, constant perception with sensors and cautioning about machines and gadgets [7, 39, 60]. The IoT-empowered observing and controlling innovations are basic in numerous business forms identified with inventory network exercises, for example, assembling, coordinations, and transportation. Observing and control frameworks gather and send key information on resource condition, hardware execution, testing, vitality utilization, and natural conditions and enable administrators and mechanized controllers to react to changes progressively anyplace, whenever. These capacities are vital to keen production network exercises where deceivability and traceability of articles are required.

Radio recurrence ID framework (RFID) is a piece of the IoT frameworks identified with inventory network exercises. RFID permits programmed recognizable proof and information catch utilizing radio waves, a tag, and a peruser. The tag can store the bigger number of information than customary scanner tags. The tag contains information as the electronic product code (EPC), a worldwide RFID-based thing distinguishing proof framework created by the auto-ID center. RFID perusers can recognize, track, and screen the items connected with labels all around, naturally, and progressively, if necessary.

Notwithstanding the RFID, remote advancements have assumed a key part in mechanical observing and control frameworks. In an average production network, observing and control applications utilize sensors, GPS, RFID labels, and sensor systems to limit scattering, robbery, misfortune, and decay in the distribution center, transportation, and store racks [26]. Sensors are utilized to keep up merchandise at the right temperature and shield them from substance spills to waste [23]. Sensor systems screen movement conditions, and route gadgets track the area of transportation vehicles to make steering more proficient. Wireless Sensor Networks (WSNs) are playing more and more a key role in several application scenarios such as healthcare, agriculture, environment monitoring, and smart metering [44]. Remote sensor arranges (WSN) comprise spatially dispersed self-sufficient sensor-prepared gadgets to screen physical or ecological conditions and can participate with RFID frameworks to better track the status of things, for example, their area, temperature, and developments [4].

WSN has been principally utilized as a part of frosty chain coordinations which includes the transportation of temperature delicate items along with an inventory network through warm and refrigerated bundling strategies [29, 63]. WSN is additionally utilized for upkeep and following frameworks. General electric conveys sensors in its fly motors, turbines, and wind ranches. By dissecting information continuously, GE spares time and expenses because of convenient preventive support. American Airline additionally utilizes sensors fit for catching 30 terabytes of information for every flight for administrations, for example, safeguard upkeep.

9.4.2 Information Sharing and Collaboration on Cloud Computing Architecture

Assets that put resources into store network connections empower data sharing and coordinated effort between inventory network accomplices [51, 54]. To accomplish SCM destinations effectively, firms need to lead a progression of SCM exercises.

Inventory network accomplices must see exact, current, and solid data from different firms with a specific end goal to encourage data sharing and joint effort [52]. To boost SCM's maximum capacity, inventory network accomplices require data sharing and coordinated effort. Data sharing and coordinated effort allude to the electronic development of data and reason-driven cooperation between individuals, among individuals and things, and between things.

Data sharing and collaboration applications improve the basic leadership of the channel accomplices through better deceivability of material stream in the production network. Between authoritative information sharing turns out to be ongoing by means of remote interchanges. As production network accomplice's increment data sharing and cooperation exercises, they lessen between firm vulnerability and reinforce vital relationship [37, 42].

Remote sensors are generally utilized for data sharing and joint effort among items and human leaders. Undertakings can enhance their abilities by utilizing the IoT-related data sharing systems inside and crosswise over enterprises [43]. Between gadget data sharing and coordinated effort, there is an advantage from the institutionalization of correspondence conventions and gadgets. Checking and control segments are frequently implanted into the data sharing and cooperation gadgets and identify a predefined occasion which triggers data sharing and joint effort among individuals and things. Late research demonstrates that there exists more potential use of IoT in data serious divisions, for example, medicinal services benefits, because of data sharing capacity of the IoT [64]. Data sharing enables gadgets and individuals to work together in the meantime, with a similar data gathered or recovered from numerous sources. Store network accomplices can share data created by IoT gadgets to track between authoritative calculated streams of materials and items and team up their conveyance exercises in view of the level of stock and the rate of offers. The data sharing and joint effort will be more gainful for worldwide production network accomplices where the vulnerability of lead time is more prominent than for the residential store network accomplices.

One key innovation of the IoT for data sharing and joint effort is fast system innovation [62]. The capability of the IoT for production network will be acknowledged by means of IoT institutionalization with which gadgets from various makers share data consistently among inventory network accomplices. Measures are required for bidirectional correspondence and data trade among things, their condition, and their advanced partners that have an enthusiasm for checking, controlling, or helping the things [5].

A cloud computing architecture will be highly suited in such data sharing environments, integrating the full extent of capabilities that cloud providers offer with internal resources to provide a comprehensive technical support and delivery solution. Cloud computing promises to enhance and optimize the technical capabilities of enterprises. Resource sharing in a multitenant infrastructure enables clients of cloud computing services who independently may not be able to make the investment for the newest technology solutions or to create the redundancy necessary to provide 100% availability. Resource constraints have caused enterprises to attempt to do more with fewer resources or to plan for redundancy only for the most critical devices, potentially leaving the enterprise vulnerable to denial-of-service attacks or at risk to outages resulting from device failure. In addition to a resilient and extensible technical architecture, cloud providers offer other benefits such as human resources, including individuals with specialized skills or personnel available on a 24/7 basis, which may otherwise have been outside the tenant's reach. Technology and people come together in an environment in which policies and processes, as well as tools and performance aids, contribute to the ability of cloud suppliers to deliver exceptional levels of service to their clients. The advantages that result from sharing technical, human, and other resources may not automatically benefit the clients/tenants. To maximize the advantage from the cloud computing systems, enterprises need to understand the resource capabilities they possess in-house as well as the resources that the external cloud services supplier makes available. Understanding what capabilities are offered and how they can be combined with internal resources, and developing a plan to leverage these combined resources, will set apart those enterprises that achieve exceptional results. To leverage both internal and cloud provider resources effectively, enterprises need to (1) understand the human and technical resource capabilities that exist in the current infrastructure and how a cloud strategy will impact the need for these or other resources; (2) define the capabilities that a cloud provider will make available as well as constraints on these resources, including periods of unavailability or priority of use; (3) consider emergency situations and resource requirements necessary to determine causes, stabilize the environment, protect sensitive and private information, and restore service levels; (4) determine how policies, practices, and processes currently support the use of technology; how transitioning to a cloud solution will require policy, practice, and process changes; and the impact these changes will have on capabilities; (5) ensure that service providers can demonstrate that personnel understand information security requirements and are capable of discharging their protection responsibilities; (6) ensure that internal staff has the skill and expertise to coordinate activities with cloud providers and that they are engaged in cloud service acquisition and ongoing management; and (7) ensure that effective channels of communication are provided with provider management and key specialists, particularly for problem identification and resolution.

9.4.3 Big Data and Data Analytics

Big data and data analytics allude to the way toward gathering, arranging, and breaking down a huge measure of information to find helpful examples and learning [19, 41] defines big data as data that cannot be handled and processed in a straightforward manner. A big dataset won't fit in memory unlike a spreadsheet data that fits easily in memory. Big data is not easy to check whether it is clean. Computations will take a long time. In big data, new data will be constantly streaming in, and the processing system must make decisions as to which part of the stream must be captured. The dataset may consist of images, natural language text, and other heterogeneous data. Also, big data dataset will probably be so large as to not fit on a single hard drive; and hence, will be stored on several different disks across a network and will be processed on a number of cores. Queries will have to be distributed and written to work across a network.

Also, frequent patterns generated over data streams also become large, which means spending a lot of time mining the patterns, and thereby it can violate one of the requirements for the data stream mining, immediate processing. In order to solve the problem, closed frequent pattern (CFP) and maximal frequent pattern (MFP) notations [55]. Huge information presents both enormous open doors and huge difficulties to our general public. The open doors incorporate better-educated business choices, more proficient store network administration and asset assignment, more compelling focusing of items and notices, and speedier turnaround of logical revelations [38]. IoT gadgets and machines with installed sensors and actuators produce a colossal measure of information to be handled by business insight and examination apparatuses for basic leadership. Various examinations concentrate on utilizing or creating compelling techniques to change over the information produced or caught by the IoT into learning helpful for basic leadership and critical thinking [8, 61]. Imaginative IoT applications are utilizing consistent extensive scale detecting, information investigation, data portrayal, and distributed computing [24].

Information created by the various accomplices' machines and gadgets are nourished into a cloud data center and converged into a production network-wide database for reproduction, information investigation, and advancement. Enormous information and information investigation applications improve the prescient upkeep of the Armada, request gauge, conveyance lead time, and client administrations. The enormous information produced by IoT has diverse attributes contrasted and general huge information due to the distinctive sorts of information gathered, of which the most traditional qualities incorporate heterogeneity, assortment, unstructured element clamor, and high repetition [13].

9.5 IoT Solutions for Streamlining the Supply Chain Activities

The IoT makes enormous open doors for the economy and people. The IoT has no limits and will affect each business and individual. With new gadgets associating into organization's frameworks, new kinds of information will be open. No less than three of the real advantages are correspondence, control, and cost investment funds which prompt changed business forms, enhanced client maintenance, new plans of action, and an expansion in the personal satisfaction. An investigation done by McKinsey Global Institute (2015) predicts that the IoT has a potential monetary effect of \$3.9 trillion to \$11.1 trillion every year in 2025 – the estimation of this effect would be proportionate to roughly 11% of the world economy in 2025.

As the Internet of Things (IoT) opens up new business openings, firms are attempting to comprehend the effects of the IoT transformation on their inventory network. Regardless of the enormous enthusiasm by network administrators to use the IoT, there is yet a lack of arrangements with the innovative part of the inventory network.

9.6 Challenges in IoT Development on Cloud Platform for Smart Supply Chain

Cloud computing represents an evolution in how enterprises acquire and use technology from external cloud service providers and interact with and support internal and external users and customers. Because of its evolutionary nature, clear guidance is needed and a direction for executive management or those responsible for determining as to how the cloud will be used and how it fits within the organizational structure and processes. Further evolution will let to better understanding, bringing more specific guidance addressing the legal, operational, and technical types of risk. Such guidance will help reduce the need to plan for building trust to achieve value and benefit at each step in the acquisition and integration of cloud solutions.

Similarly, as with any troublesome advancement, the IoT exhibits different difficulties for firms. This segment recognizes and talks about difficulties in IoT application advancement for shrewd inventory network. For instance, Gubbi et al. [24] propose that because of the blast of the information created by IoT machines, server farms will confront challenges in security, nature of administration, information investigation, stockpiling administration, server innovations, and server farm organizing. Khan et al. [34] recognize challenges from specialized points of view: naming and personality administration, interoperability and institutionalization, data protection, objects wellbeing and security, information privacy and encryption, organize security, range, and greening of IoT. Many investigations transcendently call attention to the security test of the IoT [18, 32, 57]. This segment talks about five specialized and administrative difficulties: information quality, information security, speed of innovative change, store network process upgrade, and venture avocation [40]. The talk concentrates on how these difficulties may have impacts on the advancement and reception of production network IoT applications.

9.6.1 Data Quality

Information quality is characterized by the estimation of information created and gathered by IoT objects. Information quality is essential to pick up client engagement and acknowledgment of the IoT worldview and administrations [33]. On the off chance that information is of low quality, confidence in the choice in light of the information will be low, and the reception of IoT advances can be blocked. At the point, when information gathering cost is high and the nature of the information is low, the estimation of the IoT diminishes quickly. In different circumstances, the information required for basic leadership is not accessible or not convenient. On the off chance that accomplices do not give opportune information on the sum and timing of a buy, providers would not have the capacity to build up a powerful generation design [12]. For the observing and control IoT applications, information quality is basic to the estimation of and trust in the IoT. For firms embracing the IoT applications for their inventory network, machines may not work appropriately, information gathered may not be dependable, and basic choices might be made by clients or machines in light of erroneous information. Production network accomplices must see information from different firms to be precise, current, and solid with a specific end goal to encourage data sharing and joint effort [52]. Directors would not utilize enormous information and information examination IoT applications in the event that they do not have a trust in the information quality and might not have any desire to impart their information to their accomplices.

9.6.2 Data Security

As IoT is based on the Internet, security issues of the Internet will likewise happen in IoT [32]. Once the IoT information is spilled or hindered in any piece of the IoT framework, the security of the whole IoT framework will be under threat. As the quantity of associated gadgets builds up, the quantity of the conclusion to end association focuses increments exponentially, and the quantity of potential security gaps increments in extent. Securing the IoT is an unpredictable and troublesome assignment. The number of assault vectors accessible to pernicious assailants may end up plainly amazing, as a worldwide network ("get to anybody") and openness ("get to in any case, whenever") are key principles of the IoT [57]. These security openings to singular observing and control IoT applications open the entryways for security breaks, for example, dissent of-benefit (DOS) assaults and Trojan steed assaults from programmers and other digital hoodlums. For store network cooperation, the information gathered from the IoT applications should be safely ensured to keep the privacy of inventory network accomplices' information. A current report by HP uncovers that 70% of the most regularly utilized IoT gadgets contain genuine vulnerabilities with secret word and the client have to get authorizations for it [28]. Filtering of ten most famous IoT gadgets uncovers, by and large, every gadget contained 25 powerless gaps or dangers of trading off the IoT security. The IoT gadgets normally do not utilize information encryption procedures. Feeble security will jeopardize the client protection from the reception of huge information and information examination of IoT applications for brilliant store network. Security test might be reduced via preparing application designers to join security arrangements, for example, interruption anticipation frameworks, encryptions, and firewalls into items, and urging clients to use the IoT security including those that are consolidated into their gadgets.

9.6.3 Velocity of Technological Change

The speed of mechanical change alludes to the rate of innovative change in the IoT. Development is quickening at a breathtaking rate in numerous IoT innovation wildernesses. The union of advancements in objects, systems, information stockpiling, and applications are encouraging the speed of the IoT development. The expansion in the speed of the innovative change prompts a more prominent level of natural vulnerability [6, 25]. The quick mechanical improvement in the IoT is fixing to the estimation of the market of $7.1 trillion by 2020, up from $1.9 trillion of every 2013, as the development of market request drives the innovative advances in the IoT. Remote sensor systems (WSNs) have been perceived as a promising innovation for the IoT, and mechanical advances in equipment have made hearty and practical gadgets in detecting applications [58]. Brilliant store network is being produced with propels in sensor advances and scaled down installed computational gadgets. Since innovative changes cannot be managed by an individual firm, firms need to fabricate a nearby business relationship to mutually grow new advancements and items [47]. Mendelson and Pillai [46] demonstrate that as the business condition turns out to be more questionable, firms need to refresh and overhaul existing inventory network through coordinated effort methodology all the more regularly. While different joint effort IoT applications increment the general store network efficiency, it will be harder to anticipate which IoT innovation will turn into a troublesome development in production network administration. The topic of what firms ought, to limit the danger of removal expedited by the IoT's problematic advancement stays to be replied.

9.6.4 Supply Chain Process Redesign

The IoT is an empowering influence of store network process update and in addition a problematic advancement. In the meantime, this new advancement brought out new difficulties for the outline of progressively more intricate supply chains. Changing shopper practices and the rise of transformative advances, for example, the Industry 4.0, counterfeit consciousness, huge information examination, and the Internet of Things, are driving a progression of developments with significant ramifications for the endeavors in their plan of supply chains [27]. While the utilization of the IoT in-store network is required to enhance responsibility in associations, it is indistinct how such desires will really be met in light of the fact that new advances address as well as create new responsibility requests and directors may encounter challenges in fulfilling these accountabilities [9].

Supply chain process overhaul endeavors mean to enhance arrange satisfaction, stock administration, warehousing, conveyance administration, and client relationship administration. Inventory network IoT's potential esteems talked about in the calculated structure can be completely acknowledged by coordinating it with store network process overhaul. Devaraj and Kohli [15] find that IT capital venture joined with business process reengineering enhanced gainfulness in the social insurance industry. While checking and control IoT applications may effectively meet operational proficiency, they may at present be ineffectively coordinated with joint effort concentrated production network procedures, for example, anticipating and arrange satisfaction. Subsequently, firms may make storehouses of checking and control IoT applications without completely understanding the advantages of community IoT applications. New supply process overhaul may require work investigation, work disposals, work creation, and employment preparation because of the end of existing procedure exercises and production of new process exercises. A change administration program for influenced inventory network experts should be set up before the selection of the IoT.

9.6.5 Investment Justification

While the IoT is generally new, speculation openings flourish alongside the advances of different IoT innovations. Imaginative firms are relied upon to use different capacities of the IoT advances. While actualizing an IoT domain requires enormous speculations. The venture choice where things are not open and interoperable makes it troublesome for organizations to embrace this innovation [1]. To understand the potential advantages, firms require a fitting measure with a specific end goal to legitimately survey the dangers and prizes of the innovation [41]. Numerous IoT ventures have vague issue definitions and are utilizing rising advances, along these lines, causing a higher danger of task disappointment and higher irreversibility of speculations than customary innovation ventures. When all that is said is done, interests in huge information and information examination IoT applications are more unpredictable and less secure than interests in the other two IoT application classifications. The standard measure firms commonly use to esteem ventures, net present value (NPV) and return on investment (ROI), frequently may not catch genuine estimation of the IoT extension because of the way these standard measures disregard adaptability in speculation, for example, reversibility and versatility in the assessment time frame. NPV tends to dismiss a venture's worth when there is the high vulnerability of undertaking achievement and is not reasonable for high hazard ventures, for example, ventures for huge information and information investigation. So as to esteem high hazard IoT ventures, the genuine choice approach might be more proper [41].

The benefits of cloud acquisition based on a full understanding of the cost of cloud compared with other technology platform business solutions are evaluated. The full cost of provisioning technology infrastructures and services is much larger than the cost of facilities, hardware, and software and the annual cost of human resources. The full cost of acquiring, deploying, and maintaining cloud infrastructures, platforms, or software services is more than what is represented within the cloud contract and SLA. Enterprises must also make a substantial investment in the governing structures, processes and procedures, enterprise architectures, and culture required to make technology and applications an essential element in how the business is managed and how internal and external value is assured. Investments must be made in people and their skills and in professional certifications that demonstrate their competencies. The true cost of cloud computing needs to be measured against the total investment and ongoing costs of providing similar services using internal resources. The investment in people, process, and technology creates a value structure that may not be easily replicated within the context of a cloud service strategy. To properly evaluate the costs and benefits of cloud computing, enterprises need to (1) clearly, document expected benefits in terms of rapid resource provisioning, scalability, capacity, continuity, and the cost reductions that the cloud services offer; (2) define the true life cycle cost of IT services provided internally or through a provider to have a basis for comparing expected and received value; (3) balance cost with functionality, resilience, resource utilization, and business value; (4) look beyond cost savings by considering the full benefits and value of what cloud services and support can provide; (5) periodically evaluate performance against expectations.

9.7 Conclusion

Because the IoT is such a recent development, there remains a paucity of studies on the supply chain aspects of the IoT. This makes it very challenging for enterprises to make informed decisions in regard to IoT adoption/implementation. Many enterprises invest in a variety of IoT applications for their supply chain innovation. The IoT became a reality beyond the hype. However, converting the IoT investment into real performance and profit improvement is challenging. Our chapter contributes to the literature in this emerging field by presenting a foundational knowledge for building a smart supply chain.

This chapter presented a conceptual framework of the Internet of Things (IoT) for smart supply chain management and identified three capabilities of the IoT: monitoring and control, information sharing and collaboration, and big data and data analytics. This chapter also examined how the IoT applications support the supply chain management of Fortune 200 companies. This chapter presents core and emerging IoT technologies and technology platforms for a smart supply chain. Finally, this chapter discussed five technical and managerial challenges in implementing IoT applications for smart supply chain: data quality, data security, the velocity of technological change, supply chain process redesign, and investment justification.

References

1. E. Ahmed, I. Yaqoob, A. Gani, M. Imran, M. Guizani, Internet-of-things-based smart environments: state of the art, taxonomy, and open research challenges. IEEE Wirel. Commun. **23**(5), 10–16 (2016)
2. J.S. Arlbjørn, H. de Haas, K.B. Munksgaard, Exploring supply chain innovation. Logist. Res. **3**(1), 3–18 (2011)
3. V. Arora, V. Ravi, Data mining using advanced ant colony optimization algorithm and application to bankruptcy prediction. Int. J. Inf. Syst. Soc. Chang. **4**(3), 33–56 (2013)
4. L. Atzori, A. Iera, G. Morabito, The internet of things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)
5. D. Bandyopadhyay, J. Sen, Internet of things: applications and challenges in technology and standardization. Wirel. Pers. Commun. **58**(1), 49–69 (2011)
6. M. Bensaou, Inter-organizational cooperation and the use of IT: an empirical comparison of U.S. and Japanese supplier relations. Inf. Syst. Res. **8**(2), 107–124 (1997)
7. Z. Bi, L.D. Xu, C. Wang, Internet of things for enterprise systems of modern manufacturing. IEEE Trans. Ind. Inf. **10**(2), 1537–1546 (2014)
8. S. Bin, L. Yuan, W. Xiaoyi, Research on Data Mining Models for the Internet of Things, in *Proc. International Conference on Image Analysis and Signal Processing*, (Xiamen, 2010), pp. 127–132
9. D. Boos, H. Guenter, G. Grote, K. Kinder, Controllable accountabilities: the internet of things and its challenges for organizations. Behav. Inform. Technol. **32**(5), 449–467 (2013)
10. R. Bose, Competitive intelligence process and tools for intelligence analysis. Ind. Manag. Data Syst. **108**(4), 510–528 (2008)
11. K. Butner, The smarter supply chain of the future. Strateg. Leadersh. **38**(1), 22–31 (2010)
12. S. Chapman, P. Carter, Supplier/customer inventory relationships under just-in-time. Decis. Sci. **21**(1), 35–51 (1990)
13. M. Chen, S. Mao, Y. Liu, Big data: a survey. Mobile Netw. Appl. **19**(2), 171–209 (2014)
14. H. Chesbrough, A.K. Crowther, Beyond high tech: early adopters of open innovation in other industries. R D Manag. **36**(3), 229–236 (2006)
15. S. Devaraj, R. Kohli, Information technology payoff in the health-care industry: a longitudinal study. J. Manag. Inf. Syst. **16**(4), 41–67 (2000)
16. N. Economides, E. Katsamakas, Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry. Manag. Sci. **52**(7), 1057–1071 (2006)
17. European Union, Digital agenda for Europe: monitoring and control (2015). Retrieved from http://ec.europa.eu/digital-agenda/en/monitoring-and-control
18. M.U. Farooq, M. Waseem, A. Khairi, S. Mazhar, A critical analysis on the security concerns of internet of things (IoT). Int. J. Comput. Appl. **111**(7), 1–5 (2015)
19. D. Fisher, R. DeLine, M. Czerwinski, S. Drucker, Interactions with big data analytics. Interaction **19**(3), 50–59 (2012)
20. R.K. Ganti, F. Ye, H. Lei, Mobile Crowdsensing: current state and future challenges. IEEE Commun. Mag. **49**(11), 32–39 (2011)
21. Gartner Press Release, Gartner's 2014 Hype Cycle for emerging technologies maps the journey to digital business. 2015 (2014)
22. G.L. Geerts, D.E. OLeary, A supply chain of things: the EAGLET ontology for highly visible supply chains. Decis. Support. Syst. **63**(1), 3–22 (2014)
23. M. Grunow, S. Piramuthu, RFID in highly perishable food supply chains – remaining shelf life to supplant expiry date? Int. J. Prod. Econ. **146**(2), 717–727 (2013)
24. J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, Internet of things (IoT): a vision, architectural elements, and future directions. Futur. Gener. Comput. Syst. **29**(7), 1645–1660 (2013)

25. T. Guimaraes, D. Cook, N. Natarajan, Exploring the importance of business clock-speed as a moderator for determinants of supplier network performance. Decis. Sci. **33**(4), 629–644 (2002)
26. O.K. Ha, Y.S. Song, K.Y. Chung, K.D. Lee, D. Park, Relation model describing the effects of introducing RFID in the supply chain: evidence from the food and beverage industry in South Korea. Pers. Ubiquit. Comput. **18**(3), 553–561 (2014)
27. T.S. Harrington, J.S. Srai, Designing a 'concept of operations' architecture for next-generation multi-organisational service networks. AI & Soc., (2016) https://doi.org/10.1007/s00146-016-0664-5 Published online May 17
28. HP, HP study reveals 70 percent of internet of things devices vulnerable to attack. 2015 (2014)
29. C.-F. Hsueh, M.-S. Chang, A model for intelligent transportation of perishable products. Int. J. Intell. Transp. Syst. Res. **8**(1), 36–41 (2010)
30. IBM, The Smarter supply chain of the future (2010). Retrieved January 21, 2015, from http://www-935.ibm.com/services/us/gbs/bus/html/gbs-csco-study.html
31. IDC, Worldwide and regional internet of things (IoT) 2014–2020 forecast: a virtuous circle of proven value and demand. 2015 (2014)
32. Q. Jing, A.V. Vasilakos, J. Wan, J. Lu, D. Qiu, Security of the internet of things: perspectives and challenges. Wirel. Netw **20**(8), 2481–2501 (2014)
33. A. Karkouch, H. Mousannif, H. Al Moatassime, T. Noel, Data quality in internet of things: a state-of-the-art survey. J. Netw. Comput. Appl. **73**, 57–81 (2016)
34. R. Khan, S.U. Khan, R. Zaheer, S. Khan, Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges, in *2012 10th International Conference on Frontiers of Information Technology (FIT): Proceedings*, (Institute of Electrical and Electronics Engineers Inc, Islamabad, 2012), pp. 257–260
35. D.H. Kim, J.Y. Cho, S. Kim, J. Lim, A study of developing security requirements for internet of things (IoT). Adv. Sci. Tech. **87**, 94–99 (2015). Art, culture, game, graphics, broadcasting and digital contents
36. Kimberly-Clark, Kimberly-Clark connecting with startup innovation at CES 2015 (2014)
37. D.R. Krause, R.B. Handfield, B.B. Tyler, The relationships between supplier development, commitment, social capital accumulation and performance improvement. J. Oper. Manag. **25**(2), 528–545 (2007)
38. A. Kumar, F. Niu, C. Ré, Hazy: making it easier to build and maintain big-data analytics. Commun. ACM **56**(3), 40–49 (2013)
39. M.T. Lazarescu, Design of a WSN platform for long-term environmental monitoring for IoT applications. IEEE J. Emerging Sel. Top. Circuits Syst. **3**(1), 45–54 (2013)
40. I. Lee, The internet of things (IoT) for supply chain innovation: a conceptual framework and analysis of fortune 200 companies. Asia Pac. J. Innov. Entrep. **9**(1), 81–103 (2015)
41. I. Lee, K. Lee, The internet of things (IoT): applications, investments, and challenges for enterprises. Bus. Horiz. **58**(4), 431–440 (2015)
42. J. Lee, Y. Kim, Effect of partnership quality on IS outsourcing success: conceptual framework and empirical validation. J. Manag. Inf. Syst. **15**(4), 29–61 (1999)
43. F. Al-Turjman, S. Alturjman, Context-sensitive access in industrial internet of things (IIoT) healthcare applications. IEEE Trans Ind Inf **14**(6), 2736–2744 (2018)
44. L. Mainetti, L. Patrono, A. Vilei, Evolution of wireless sensor networks towards the internet of things: a survey. 2011 19th international conference on Software, Telecommunications and Computer Networks (SoftCOM), 1–6 (2011)
45. J. Mariani, E. Quasney, M.E. Raynor, Forging links into loops: the internet of things' potential to recast supply chain management. Deloitte Review, 17 (2015)
46. H. Mendelson, R. Pillai, Clockspeed and information response: evidence from the information technology sector. Inf. Syst. Res. **9**(4), 415–433 (1998)
47. J.T. Mentzer, S. Min, Z.G. Zacharia, The nature of interfirm partnering in supply chain management. J. Retail. **76**(4), 549–568 (2000)
48. R.E. Miles, C.C. Snow, Organization theory and supply chain management: an evolving research perspective. J. Oper. Manag. **25**(2), 459–463 (2007)

49. J. Mineraud, O. Mazhelis, X. Su, S. Tarkoma, A gap analysis of internet-of-things platforms. Comput. Commun. **89-90**, 5–16 (2016)
50. F. Al-Turjman, S. Alturjman, 5G/IoT-enabled UAVs for multimedia delivery in industry-oriented applications. Multimed Tools Appl J (2018). https://doi.org/10.1007/s11042-018-6288-7
51. R. Patnayakuni, A. Rai, N. Seth, Relational antecedents of information flow integration for supply chain coordination. J. Manag. Inf. Syst. **23**(1), 13–49 (2006). https://doi.org/10.2753/MIS0742-1222230101
52. K.J. Petersen, G.L. Ragatz, R.M. Monczka, An examination of collaborative planning effectiveness and supply chain performance. J. Supply Chain Manag. **41**(2), 14–25 (2005)
53. L. Ping, Q. Liu, Z. Zhou, H. Wang, Agile Supply Chain Management over the Internet of Things, in *2011 International Conference on Management and Service Science (MASS)*, (Wuhan, 2011), pp. 1–4
54. K. Pramatari, Collaborative supply chain practices and evolving technological approaches. Supply Chain Manag. **12**(3), 210–220 (2007)
55. R.V. Priya, A. Vadivel, S. Thakur, Maximal pattern mining using fast CP-tree for knowledge discovery. Int. J. Inf. Syst. Soc. Chang. **3**(1), 56–74 (2012)
56. P.J. Reaidy, A. Gunasekaran, A. Spalanzani, Bottom-up approach based on internet of things for order fulfillment in a collaborative warehousing environment. Int. J. Prod. Econ. **159**, 29–40 (2015)
57. R. Roman, J. Zhou, J. Lopez, On the features and challenges of security and privacy in distributed internet of things. Comput. Netw. **57**(10), 2266–2279 (2013)
58. D. Singh, G. Tripathi, A.J. Jara, A survey of internet-of-things: future vision, architecture, challenges and services. 2014 IEEE world forum on internet of things (WF-IoT) (2014)
59. H. Sundmaeker, P. Guillemin, P. Friess, S. Woelfflé, Vision and challenges for realising the internet of things, CERP-IoT – Cluster of European Research Projects on the Internet of Things. 2014 (2010)
60. F. Tao, Y. Zuo, L.D. Xu, L. Zhang, IoT based intelligent perception and access of manufacturing resource towards cloud manufacturing. IEEE Transactions on Industrial Informatics **10**(2), 1547–1557 (2014)
61. C.-W. Tsai, C.-F. Lai, M.-C. Chiang, L.T. Yang, Data mining for internet of things: a survey. IEEE Commun. Surv. Tutorials **16**(1), 77–97 (2014)
62. H. Wang, T. Zhang, Y. Quan, R. Dong, Research on the framework of the environmental internet of things. Int J Sust Dev World **20**(3), 199–204 (2013)
63. C.C. White III, T. Cheong, In-transit perishable product inspection. Res. E Logist. Transp. Rev. **48**(1), 310–330 (2012)
64. B. Xu, L.D. Xu, H. Cai, C. Xie, J. Hu, F. Bu, Ubiquitous data accessing method in IoT-based information system for emergency medical services. IEEE Trans. Ind. Inf. **10**(2), 1578–1586 (2014)
65. M. Porter, The value chain and competitive advantage, Chapter 2 in competitive advantage: creating and sustaining superior performance (Free Press, New York, 1985), pp. 33–61
66. Z. Zhou, Z. Zhou, Application of internet of things in agriculture products supplies chain management. International Conference on Control Engineering and Communication Technology (ICCECT) (2012)
67. L.J. Krajewski, L.P. Ritzman, M.K. Malhotra, *Operations Management: Processes and Supply Chains*, 10th edn. (Pearson, 2013)
68. C. Brian Gibson, C. Defee, H. Chen, J.B. Hanna, *The Definitive Guide to Integrated Supply Chain Management: Optimize the Interaction between Supply Chain Processes, Tools, and Technologies (Council of Supply Chain Management Professionals)*, 1st edn. (Prentice Hall Inc, 2014)
69. Miguel Gastón Cedillo-Campos, *Supply chain clustering: The next logistics frontier?*, International Congress on Logistics & Supply Chain, CiLOG2014

Chapter 10
A Conceptual Framework for Security and Privacy in Edge Computing

Bragadeesh S. A. and Umamakeswari Arumugam

10.1 Introduction

Ubiquitous connectivity is the need of the hour with the ever-increasing volume of devices getting access to the Internet. Several key technologies drive the proliferation of IoT solutions across multifaceted application domains that are extensively discussed, researched, and adopted. The key challenge in IoT is processing the data from devices and performing meaningful interpretation to derive insights and take decisions.

Cloud computing has made possible the widespread adoption and proliferation of IoT. It has been the backbone of IoT deployments and applications which has helped in achieving primary network objectives like storage, communication, and computation. It also offers services that can be tailored according to user and application requirements. Cloud computing consists of a pool of resources, typically consisting of services like networking, storage, and computing, which are organized to cater the needs of multiple end users based on a multitenant model. The access to these resources is via standard mechanisms and is usually available across a network. The varieties of service offered by the cloud computing paradigm include SaaS (Software-as-a-Service) in which applications are offered as service, IaaS (Infrastructure-as-a-Service) which offers computing resources as a service, and finally PaaS (Platform-as-a-Service) which provides software and hardware tools for development. The widely adopted deployment models of cloud computing include public and private clouds in which organizations deploy cloud services to any end user or within their own cloud computing platform, respectively. Cloud computing is capable of providing the following value additions – elasticity, ubiquity, reduced

Bragadeesh S. A. (✉) · U. Arumugam
School of Computing, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3_10

management effort, pay-as-you-use, and convenience which has been a driving force for the creation of a rapidly growing industry across the globe which has its worth in billions [9].

Regardless of all these benefits, there are few areas in which cloud computing does fall short of meeting the application requirements, which include increased network latency, jitter, lack of ability to access local contextual information, and lack of support for mobility of users. For an application that is delay sensitive, these requirements must be met. Due to these reasons, there has been an evolution of new computing paradigms such as fog computing, mobile cloud computing, and mobile edge computing, among others [2, 7]. Edge computing leverages the resources available on the local edge to meet the specific application requirements mentioned. Edge computing can bring the services offered by cloud computing to the network edge and the essential functionalities like communication, storage, and computation closer to the devices and users.

The primary objective of these computing technologies is to bring about the capabilities of cloud computing closer to the user or near the network edge. Normally most edge computing paradigms follow the structure as shown in Fig. 10.1. The edge layer devices as specified can be microdata centers, gateways, and dedicated

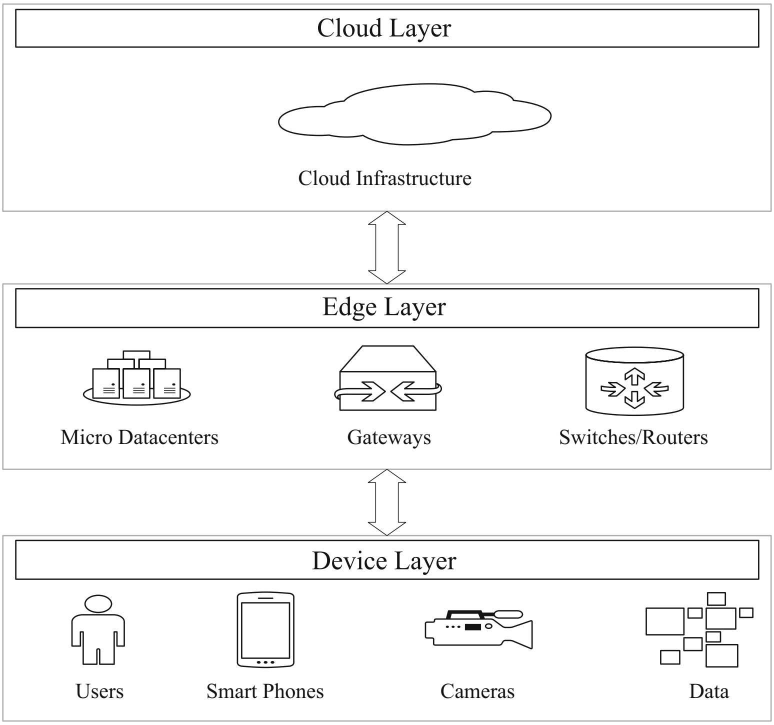


Fig. 10.1 Edge computing model

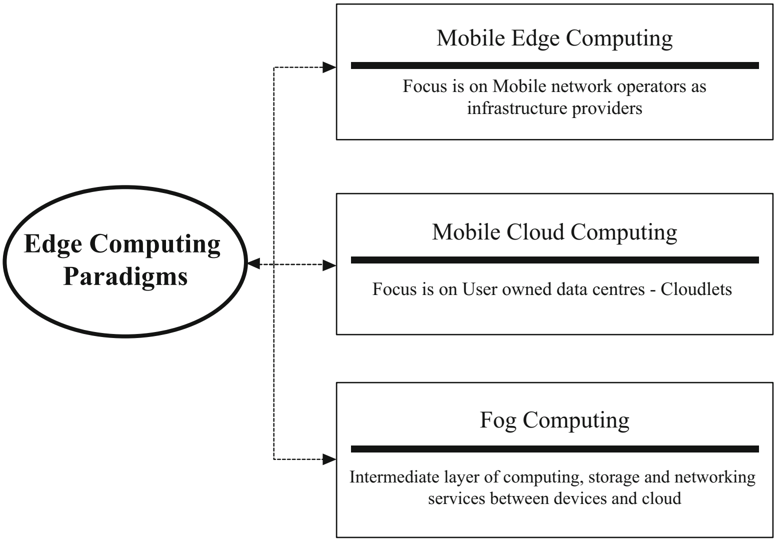


Fig. 10.2 Edge computing technology paradigms

routers/switches. Microdata centers are capable of functioning autonomously and collaborate with other such devices to offer services to end users, to third-party clients, and sometimes to the infrastructure providers too. They are still connected to the traditional cloud infrastructure which creates an opportunity to establish a hierarchical architecture to provide platforms for the management and registration of user services. An edge infrastructure which is owned by a single service provider referred to as trust domain can collaborate with other such trust domains, forming an environment where it is possible to serve multitude end users.

Figure 10.2 is helpful in pointing out the key functional differences in the different edge paradigms. A detailed comparison between the various edge paradigms is available in Sect. 10.2. There are several similarities among them. The architectures, services, mechanisms, and protocols applicable for one of them are applicable and can be adopted for the other.

The security and privacy mechanisms for edge computing paradigms are still in its emerging stage, and given the growing importance of the field, it is critical to identify the possible threats. It may not be possible to reciprocate the currently available tools and methods for security and privacy in cloud computing for edge computing owing to its specific attributes such as heterogeneity, low latency, location awareness, spatial distribution, and mobility. Even though some of these issues can be addressed using existing security schemes, new issues arise due to the above-mentioned special characteristics of edge computing. It might be even possible to adopt and analyze the existing security and privacy mechanisms that are implemented for other enabling technologies and related computing paradigms. Some of the enabling technologies which play a significant role in realization of edge computing include virtualization technologies, wireless networks, peer-to-peer and distributed networks, and software-defined networking (SDN), among others [1, 2].

The rest of the content is organized as follows. Section 10.2 elucidates the similarities and differences between the various edge computing paradigms. Section 10.3 introduces the basic components of security and privacy for edge computing and summarizes the various threats that target the edge paradigm. Section 10.4 presents a framework for implementing security and privacy for edge paradigm. Section 10.5 highlights the various challenges involved in implementation of security and privacy solutions. Section 10.6 provides the concluding remarks and briefly talks about the future directions. It is our belief that this paper provides better clarity for researchers and users regarding the concepts of edge computing and tries to stress the importance of ensuring the security and privacy as a key requirement for application and system developers.

10.2 Similarities and Differences Between Edge Paradigms

Table 10.1 gives an overall idea of how the edge paradigms differ from the traditional cloud structure. This comparison is necessary to understand the challenges that arise when we need to implement solutions in edge computing. It is evident that all the edge paradigms have the common goal of bringing the capabilities of cloud computing closer to the edge. They all provide in some way the multitenant virtualization infrastructure which helps in location-specific provisioning capabilities and access to nearby computational resources when required. We have taken into account MEC, MCC, and fog computing for our comparison [1, 5]. As mentioned earlier all of them have own similarities among them and uniqueness as well. They

Table 10.1 Differences between edge computing and cloud computing

Features	Cloud computing	Edge computing
Computational capacity	High	Medium to low
Size and operating mode	Centralized large servers	Distributed smaller servers
Applications	Delay tolerant, computationally intensive	Low latency, real-time operation, high QoS
Fronthaul/backhaul communication overhead	High	Low
Mobility support	Low	High
Management	Service provider	Local business
Deployment	Requires complicated deployment planning	Ad hoc deployment/minimal or no planning
User device	Computers, limited mobile devices	Mobile smart wearable devices
Network access type	WAN	LAN/WLAN

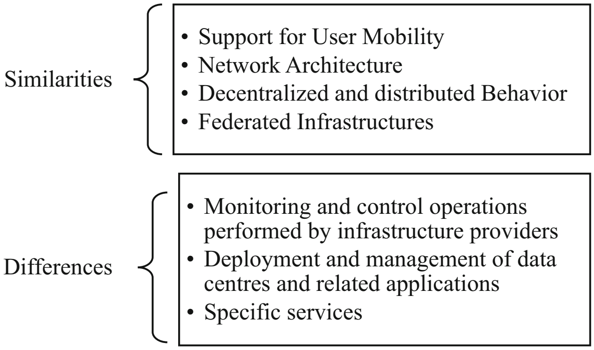


Fig. 10.3 Similarities and differences between edge paradigms

provide a similar set of benefits that are majorly due to the closeness of the edge data centers. The benefits are as follows:

- Low and predictable jitter and network latency
- Location-based context-aware information
- Support for scalability
- High availability of services

Figure 10.3 summarizes the similarities and differences between the edge paradigms. The predominant similarity is that all these techniques basically can provide support for mobility. They have consideration for device mobility and have specific management mechanisms that are located at the application level itself or use virtual instances of devices to accommodate for the same. The next similarity is the overall architecture which has support to enable the edge paradigms to behave as an extension of cloud infrastructure. The network elements act in a distributed and decentralized way capable of provisioning services and taking decisions autonomously. They can also cooperate with each other to reduce the dependency on the central cloud setup. The different edge data centers can exchange information among them and coexist to form a federated infrastructure.

Moving on to the differences, even though the edge paradigms have a common objective, they have few basic differences in the way they fulfill the objective. MEC deployments primarily focus on realization of 5G, whereas fog nodes can provide services to other applications which have their own servers, gateways, access points, and so on. In case of MCC, it is highly distributed in nature, and device instances can perform their own service provisioning. This difference in management and deployment of the data centers determines who will be service providers. Another difference is that the applications supported are determined by the choice of service providers. MEC provides support for operators to work closely with other third-party service providers, which enables for thorough testing and possibility for customized integration. This applies true even for fog computing also. The final difference is that MCC paradigm provides specific services that are not related to virtualization but enables support for execution mechanisms that are distributed in nature. This is highly beneficial for devices that have severe resource constraints.

Although there are differences, still the security and privacy techniques used for one paradigm can be easily adopted for the other.

10.3 Overview of Security, Privacy, and Threats in Edge Computing

As emphasized earlier, edge computing is not a one-to-one replacement of cloud computing, whereas it helps in augmenting and broadening the services of cloud computing. The principal services offered by edge computing include storage, computation, data sharing, and networking [12]. With respect to these four services, unique security and privacy requirements emerge as shown in Fig. 10.4.

10.3.1 Security in Edge Computing

The key attributes whenever we are trying to implement security solutions are authentication, integrity, confidentiality, trust, access control, data security, and privacy [1, 11]. The various components of providing end-to-end security and privacy are as follows.

10.3.1.1 Network Security

The configurations performed by the network administrator and the network management information must be secured and isolated from the normal data flow.

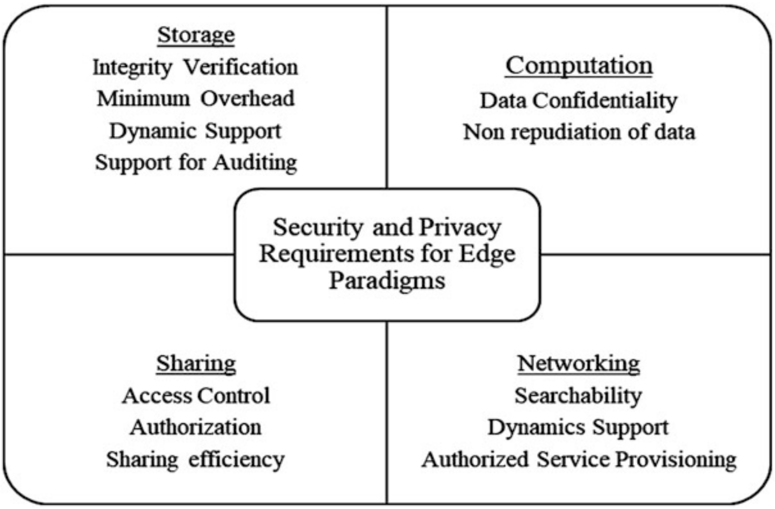


Fig. 10.4 Summary of security and privacy requirements for edge computing

This is required since the edge devices are distributed across the network and the cost of maintenance is high. However, a new network concept called software-defined networking (SDN) can come to the rescue. The benefits of SDN include but not limited to reducing the burden of management and implementation, improving the scalability of the network, decreasing the cost of maintenance, providing prioritization and isolation of network traffic, and enhancing the access control of network resources, cooperation, and network sharing.

10.3.1.2 Data Storage Security

Like cloud computing, the data storage in edge platform is also outsourced. It is tough to ensure data integrity since there are a lot of possibilities for data loss and data modifications. If the data present in one device is compromised, the attackers could easily abuse the data stored to fulfill their own needs. Hence, provision for data storage auditing should be done. Third-party auditing services are available, which are provided by the infrastructure providers, and the application users should be aware of the auditing policies. It is possible to use encryption techniques to ensure integrity, verifiability, and confidentiality of the data to check for untrusted network entities and allow the user to check the stored data. Since all data need not be essentially present on all the available storage resources, reduced latency and support for dynamic access are required.

10.3.1.3 Data Computation Security

The data computation performed at the edge servers and devices should be secured and verifiable. The security of computations can be ensured by using data encryption techniques, which prevents data visibility to any hackers/attackers. Since the microdata centers have the provisions to off-load some of the computations to other data centers, a mechanism to verify the computation results and establish trust between the two entities becomes a necessity. The computational accuracy should be verifiable by the user. Use of data encryption for the data that is being forwarded to the edge data centers from end devices and from one data center to another helps in ensuring data integrity and protection. One of the popular mechanisms to verify the data security is using data search technique in which a keyword-based search is performed in the encrypted data repositories. This enables users to securely search for user data amidst encrypted data. This helps in maintaining the secrecy of encryption.

10.3.1.4 End Devices Security

Since the end devices are typically less powerful and have limited access to the surroundings, they can be easily tampered with. Any hacker may try to take control of a device and make it a rogue node to retrieve essential network management data. They can hinder the normal behavior by corrupting the device data and increase the frequency of data access by sending fake information. They can also use the end device to propagate false data across the network to potentially create discrepancies and disrupt the normal network behavior. Hence, proper efforts should be made for ensuring the security of the end device.

10.3.1.5 Access Control

Enforcement of access control mechanisms can provide dual benefits of security and privacy. Since the edge platform is distributed and decentralized in nature, a good access control policy acts as a defensive shield to mitigate unauthorized device and service access. Access control helps in realizing the interoperability and collaboration among microdata centers that are provided by different service providers and are separated across different geo-locations. A robust access control mechanism is required to meet the design goals, accommodate mobility, low latency, and interoperability.

10.3.1.6 Intrusion Detection

Intrusion detection techniques help in identifying malicious data entries and detect device anomalies. They can be used to carefully investigate and analyze the behavior of devices in the network and provide methods to perform packet inspection, which helps in early detection of denial-of-service attacks, integrity attacks, and data flooding, among others. The primary challenge when implementing intrusion detection for edge platform is to accommodate for scalability, mobility, and low-latency requirements.

10.3.2 Privacy in Edge Computing

Privacy in edge computing can be enforced in a threefold manner: user privacy, data privacy, and location privacy. User privacy deals with protecting the privacy of user's credentials and their frequency of accessing the data. A potential hacker may observe the usage pattern and try to pretend to be legitimate user and access the necessary information. Hence, ensuring the privacy of user is required. Data privacy-preserving techniques may be available at the edge and cloud level, but making them available at the resource-constrained end device is a design challenge. This helps in avoiding unauthorized access of the network by hackers and helps in preserving the integrity of the network.

Since the edge data centers may be spatially distributed, ensuring location privacy is also a prime requirement. Since the location data can help in deciphering the environment-related information, the choice of edge data centers has to be carefully planned. Identity obfuscation is a technique used to protect the identity of an edge device from the user even though the user is nearby. When more than one edge nodes are being used by an application, it is relatively easier for an attacker to isolate the location, since the proximity of edge nodes gives them an idea that the end device should be in between the edge nodes. Ensuring location privacy without affecting the computational overhead is of primary importance.

10.3.3 Nature of Threats in Edge Computing

Any possible attempt to disrupt the normal functioning of a network or a system is deemed as a threat. Only when we understand the nature of threats, it is possible to define security solutions for it. Hence, the basic knowledge about threats and how they can be mitigated is essential. All the possible threats in edge computing environment have been presented in Table 10.2.

Table 10.2 List of threats

Components	Possible threats
Data	Data replication
	Data manipulation
	Data deletion
	Illegal data access
	Eavesdropping
Network	Denial of service
	Man-in-the-middle
	Physical damage
	Service manipulation
	Rogue server
	Privacy leakage
	Jamming
	Congestion
	Black hole
Communication	Data loss
	Data breach
	Sniffing
	Message replay
	Impersonation
Services	Service manipulation
	Rogue infrastructure
	Privacy leakage
	Insecure APIs
Devices	Physical damage
	Misuse of resources
	Flooding
	Power failure

The threats can be grouped according to the component of the system they try to attack [6]. The implementation of solutions to these attacks can be made possible by a combination of security attributes such as authentication, confidentiality, integrity, availability, and privacy.

10.4 Framework for Security and Privacy in Edge Computing

Figure 10.5 represents the basic components of the proposed conceptual framework. The framework consists of a hierarchical structure in which the communication and exchange of data can be done using APIs. There are three levels: device management, security services, and application and infrastructure. The proposed framework aims to bring into account the necessary security and privacy requirements which are specific to edge computing platform. The framework can be implemented in an edge

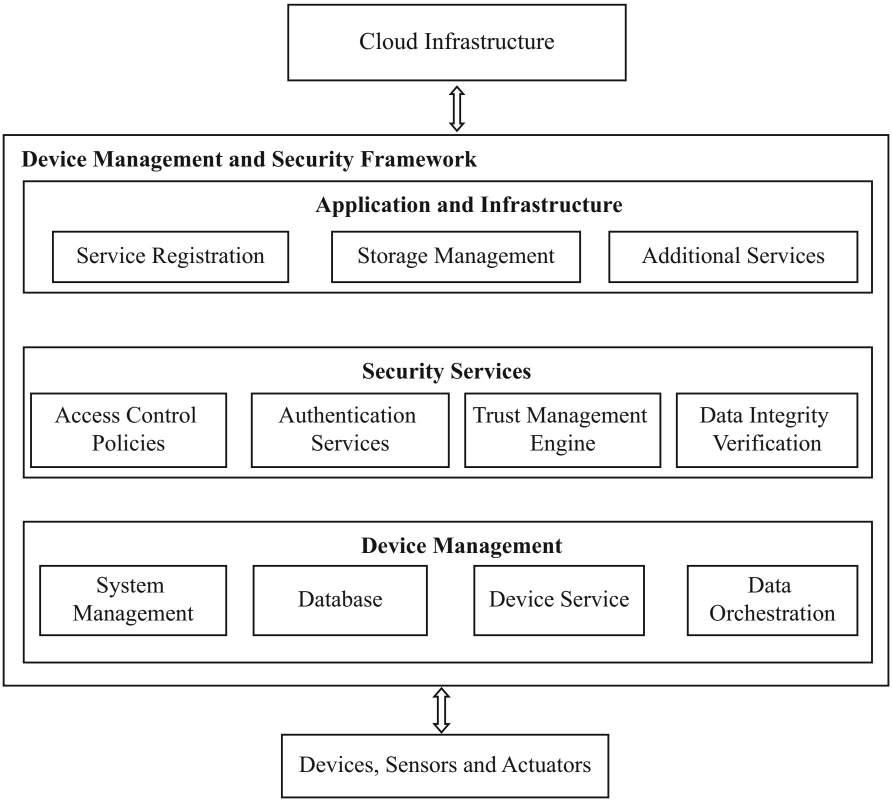


Fig. 10.5 Security and privacy framework for edge computing

gateway or a microdata center. The suitability of the framework for implementation in the resource-constrained end devices needs further investigation.

10.4.1 Device Management

The primary functionalities of the device management layer are to manage the end devices, create database, manage the services provide to the end device, and perform data orchestration operations. The status of the end devices, access control, maintaining integrity of the data stored, and providing the necessary data from service providers are all taken care by the device management layer. Network function virtualization (NFV) is one of the popular and emerging technologies which can perform the tasks mentioned above. NFV in convergence with SDN can help in realizing the full potential of edge computing platform. NFV provides virtualization of infrastructure services, orchestration, and management functionalities.

10.4.2 Security Services

The security services layer is the heart of the framework which performs the critical tasks such as creating, updating, and maintaining the access control policies. It also ensures authentication of devices, service providers, and the data that is communicated. It also makes use of a trust management engine which is used for accomplishing two objectives. It governs the privacy of the overall system and generates rules for implementing intrusion detection mechanism. The data integrity verification component performs the verification of data across the system and computations done by other edge devices. Establishment of well-defined APIs and integration of the modules with the application and device management layer help in improving the overall performance of the framework and achieving the edge computing specific requirements.

10.4.3 Application and Infrastructure

This layer is responsible for keeping track of the network-wide application interactions and governs the authentication policies for accessing the network devices. The service provisioning details of neighboring data centers and making the decision of whether cloud support is required or not are carried out by this layer. Also, the services offered by the infrastructure providers are monitored and provisioned for access to the end devices. This layer is significant because edge platform caters to multitenant application services, hence implementing security measures for it is an essential task.

10.5 Security and Privacy Challenges in Edge Computing

Edge computing has opened new arenas and created new opportunities to build solutions that can augment existing cloud-based solutions. In order to explore the full potential and bring about the widespread adoption of edge computing solutions, one key area of focus is how to secure the system components and how to ensure the privacy of user and organization data [6]. Figure 10.6 points out some of the major challenges that need to be addressed [3, 4, 8, 10]. The challenges can be summarized as below.

- The added security mechanism should not increase the computation overhead and hamper the storage space required for other system operations.
- Edge computing mechanisms use cache management techniques which are prone to side channel attacks; hence, care should be taken to prevent leakage of private and sensitive data.
- For applications that stream data continuously for monitoring and management purposes, hence due to increased number of devices, the volume data to be processed is typically large. To detect the anomaly, packet filtering mechanisms need to be implemented which might require additional memory and processing power.
- It is difficult to manage user identity, scalability, monitoring, performance, data security, and considering threats from insiders due to the large number of users who share the resources and application in the edge computing environment.

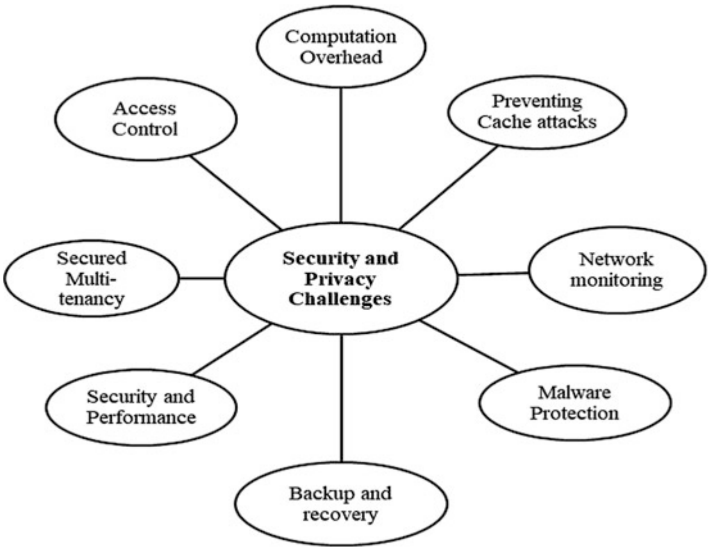


Fig. 10.6 Security and privacy challenges in edge computing

- The trade-off between application performance and security is solely dependent on the user or application's requirements; hence, establishing clear performance benchmarks without compromising security is important.
- Depending on the application and user requirements, the edge platform may have high data throughput and relatively lesser amount of data stored which may lead to problems for data storage and recovery.
- By using virtualization services, edge computing provides multitenant application access leads to exposing APIs to multiple users which may lead to insider attacks and data breaches.
- Due to the increased number of devices, implementing a common access control policy becomes a tedious task. Accounting for mobility of devices is also important.
- To ensure the privacy of user's data, the data privacy policies of the edge platform as well as the cloud infrastructure must be in sync which might lead to potential challenges. The privacy policies of different infrastructure providers may not be the same.

10.6 Conclusion

In this work, we have defined the various edge paradigms and listed out their similarities and differences. We have also made a comparison of edge computing with cloud computing to drive in the importance of new implementation challenges and approaches required for edge computing solutions. We have analyzed the various aspects of security and privacy for edge computing and also brought into light some of the possible threats. A conceptual framework for ensuring security and privacy in edge computing has been discussed along with various components involved in it. The reality to be considered here is that security and privacy solutions for edge computing are still in their nascent stages, and, hence, we have analyzed the possible challenges that might arise. This work is intended to be an eye-opener for developers and technology enthusiasts who are involved in developing security and privacy solutions for edge computing.

The future scope of this work is to develop a system which consists of a knowledge base, rule-based inference engines, and a decision support toolbox that could be added to an edge computing platform which helps in mitigating possible security attacks and ensures the privacy of user data.

Acknowledgment The authors wish to express their sincere thanks to the Department of Science & Technology, New Delhi, India (Project ID: SR/FST/ETI-371/2014), and express their sincere thanks to the INSPIRE fellowship (DST/INSPIRE Fellowship/2016/IF160837) for their financial support. The authors also thank SASTRA Deemed to be University, Thanjavur, India, for extending the infrastructural support to carry out this work.

References

1. N. Abbas et al., Mobile edge computing: a survey. IEEE IOT J. **5**(1), 450–465 (2018)
2. Y. Ai, M. Peng, K. Zhang, Edge cloud computing technologies for internet of things: a primer. Digit. Commun. Netw. **4**(2), 77–86 (2018)
3. F.A. Alaba et al., Internet of things security: a survey. J. Netw. Comput. Appl. **88**(April), 10–28 (2017)
4. A. Alrawais et al., Fog computing for the internet of things: security and privacy issues. IEEE Internet Comput. **21**(2), 34–42 (2017)
5. A.V. Dastjerdi, R. Buyya, Fog computing: helping the internet of things realize its potential. Computer **49**(8), 112–116 (2016)
6. S. Khan, S. Parkinson, Y. Qin, Fog computing security: a review of current applications and security solutions. J. Cloud Comput. **6**(1), 19 (2017)
7. P.G. Lopez et al., Edge-centric computing: vision and challenges. ACM SIGCOMM Comput. Commun. Rev. **45**(5), 37–42 (2015)
8. B.A. Martin et al., OpenFog Security Requirements and Approaches, (November, 2017), pp. 1–6
9. R. Roman, J. Lopez, M. Mambo, Mobile edge computing, Fog et al.: a survey and analysis of security threats and challenges. Futur. Gener. Comput. Syst. **78**, 680–698 (2016)
10. J. Tan, R. Gandhi, P. Narasimhan, *Challenges in Security and Privacy for Mobile Edge-Clouds* (Parallel Data Laboratory, Carnegie Mellon University, 2013). http://www.pdl.cmu.edu/PDL-FTP/associated/CMU-PDL-13-113.pdf
11. S. Yi, Z. Qin, Q. Li, Security and privacy issues of fog computing: A survey, in *10th International Conference on Wireless Algorithms, Systems, and Applications*, ed. by K. Xu, H. Zhu (Eds), (Springer International Publishing, 2015), pp. 685–695. https://doi.org/10.1007/978-3-319-21837-3_67
12. Y. Guan, J. Shao, G. Wei, M. Xie, Data security and privacy in fog computing. IEEE Netw. **99**, 1–6 (2018)

Index

A

Africa, mist computing challenges
  assimilation and domestication, 125
  business blueprint, 126
  drivers supporting growth, 126–127
  failure, faults and risk, 125–126
  funding opportunities, 126
  inclusiveness and governance, 125
  skills gap and digital divide, 124–125
  technical recommendations, 128–129
  technology transfer, 125

B

Big data
  IoT applications, 161–162
  mist and dew computing, 111
  in smart cities, SDN, 92–93
Bioacoustic sensor, 135–138
Blockchain technology, 115, 121, 123
Bowel activity monitoring
  battery charger circuit, 137
  bioacoustic sensor, 135–137
  bowel sounds, 135
  electronic stethoscopes, 134
  high-pass filter and normalization, 139
  intestinal activity detection, 141–142
  intestinal sounds, 135
  IoT-driven sensor device, 136
  LMS algorithm, 134
  operational amplifier circuit, 138
  power spectrogram, 142
  regulator circuit scheme, 137
  signal processing steps, 137–141

C

Cloud-based radio access network (CRAN), 33, 37
Cloud computing
  deployment clouds, 52
  vs. edge computing, 176–178
  mist and dew computing, 111
  service models, 52–53
  supply chain management
    data quality, 164
    data security, 164–165
    investment justification, 166–167
    redesign process, 165–166
    technological change velocity, 165
Cloud provider
  cloud service brokers, 11
  future improvement, 12
  Internet of Things
    black box syndrome, 10
    communication models, 9
    interconnectivity, 10
    notebook PC *vs.* tablet, 11
    privacy considerations and legal issues, 10
    smart objects, 9
    volume, 9–10
  pattern prediction technique, 8
  resource allocation
    Amazon's charging models, 3
    instance type, 4
    power consumption, 6–7
    revenue, 5
Cloud service brokers (CSBs), 11
Communications service providers (CSPs), 35

© Springer Nature Switzerland AG 2019

F. Al-Turjman (ed.), *Edge Computing*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-319-99061-3

Community cloud, 52
Convolutional neural networks (CNNs), 43
Council of Supply Chain Management Professionals (CSCMP), 148

D
Data reduction techniques
  comparsion, 65
  compression, 62
  data aggregation, 62–63
  data handling solutions, 61
  data mining and machine learning, 63
  data pre-processing, 63
  data redundancy, 63
  data sampling, 61–62
  FlexAnalytics, 61
  network theory, 63
  perceptually important points, 61
Deep learning, MEC
  automatic detection process, 43–45
  base station routers, 44
  machine learning techniques, 44
  mobile sensing, 43
  security threats, 45
Deep neural networks (DNNs), 43
Domain name system (DNS), 39
Dynamic voltage and frequency scaling (DVFS), 6

E
Edge application, 20–21
Edge communication, 19–20
Edge computing, 17–19
  application and infrastructure, 183
  challenges and solutions
    data abstraction, 65
    naming system, 65
    privacy and security, 66, 184–185
    programmability, 64
    service management, 66
  cloud hierarchy, 19
  cloudlet, 57
  communication network, 26–30
  concept, 55
  data consumer to producer, 56
  device management, 183
  fog computing, 59–60
  integer linear programming formulation, 22–24
  IoT-cloud infrastructure, 18
  IoT technologies, 73
  mist and dew computing, 117, 118
    architecture, 116
    characterstics, 117
    cloudlets, 114
    fog computing, 112–113
    microdata centre, 114
    mobile edge computing, 113
    M2M, 115
  mobile edge computing, 57–59
  nature of threats, 181–182
  optimized resource allocation, 24–26
  privacy mechanisms, 180–181
  pull from IoT, 56
  push from cloud services, 55–56
  QoS requirements, 54
  security mechanisms
    access control, 180
    data computation, 179
    data storage, 179
    end devices, 179–180
    intrusion detection, 180
    network security, 178–179
    services layer, 183
  technology paradigms, 174, 175
  time sensitivity, 18
European Telecommunications Standards Institute (ETSI), 34
Everything-as-a-Service (EaaS), 52

G
Gastrointestinal track motility monitor (GTMM), 134
Generative adversarial networks (GANs), 43
Google Cloud Platform (GCP), 79

H
Hardware-as-a-Service (HaaS), 52
Hybrid cloud, 52

I
Industry Specifications Group (ISG), 34
Infrastructure-as-a-Service (IaaS), 52
Internet of Things (IoT), 2, 53
  biosensors, 85
  cloud platform, SCM
    data quality, 164
    data security, 164–165
    investment justification, 166–167
    redesign process, 165–166
    technological change velocity, 165
  connected cars, 79

construction industry, 81–82
data processing layer stack, 73
driven sensor device, 136
energy management, 83
environmental monitoring, 84
healthcare systems, 85–86
Industrial IoT, 72, 75–76
infrastructure management, 74
mist and dew computing, 109–111
peer-to-peer networking, 73
poultry farming, 87
precision agriculture, 86–87
self-parking vehicles, 79–80
smart cities, 81
smart homes, 82
smart panels, 82–83
smart retail, 76–77
smart supply chain, 77–78
smart surveillance, 87
smart wearables, 84
smoke detectors, 80–81
supply chain management
big data and data analytics, 161–162
cost reduction, 154
data sharing and collaboration, 159–161
innovation stage, 156
logistics, 155
monitoring and control, 158–159
network architecture, 157
open advancement, 157
optimal distribution network, 154
RFID labels, 153, 154
streamline solutions, 162–163
Things as a Service, 158
technology roadmap, 72
transportation and logistics, 78

L
Low-power wide area network (LPWAN) controls, 64

M
Machine-to-machine communication (M2M), 54, 115
Microdata centre (MDC), 114
Mist computing
blockchain technology, 121
challenges in Africa
assimilation and domestication, 125
business blueprint, 126
drivers supporting growth, 126–127
failure, faults and risk, 125–126
funding opportunities, 126
inclusiveness and governance, 125
skills gap and digital divide, 124–125
technical recommendations, 128–129
technology transfer, 125
comparsion, 119
and dew computing
big data, 111
blockchain technology, 115
cloud computing, 111
edge computing, 112–114
fluid computing, 114
Google search trend, 109
IoT technology, 109–111
Industry 4.0
accurate predictions and better decisions, 124
blockchain technology, 123
cost reduction, 123–124
cross-continental operations, 122
direct and impactful value creation, 124
MindSphere, 123
predictive maintenance, 123
secure, safer and sustainable operation, 124
uptime improvement and availability, 124
smart cities, 119–120
Mobile edge computing (MEC), 34
applications, 39–41
architecture, 36, 37
challenges, 41–42
cloud-RAN, 37
communications service providers, 35
deep learning
automatic detection process, 43–45
base station routers, 44
machine learning techniques, 44
mobile sensing, 43
security threats, 45
domain name system, 39
edge computing, 57–59
ETSI, 34
framework, 36–39
ISG, 34
mobile edge orchestrator, 38
radio access network, 36
radio applications cloud servers, 34

N
Named data network (NDN), 101

P

Platform-as-a-Service (PaaS), 52
Private cloud, 52
Proposed system architecture, SDIoT, 93–94
  application level, 100–101
  computational methods, 96
  data collection and intermediate levels, 95–96
  data management and processing level, 96–97
  HDFS system, 100
  message queue, 97–100
  named data network, 101
  threshold limit lalue, 97
Public cloud, 52

R

Radio access network (RAN), 36
Radio applications cloud servers (RACS), 34
Radio network controller (RNC), 36
Recurrent neural networks (RNNs), 43
RedEdge model, 64

S

SDIoT, proposed system architecture, 93–94
  application level, 100–101
  computational methods, 96
  data collection and intermediate levels, 95–96
  data management and processing level, 96–97
  HDFS system, 100
  message queue, 97–100
  named data network, 101
  threshold limit lalue, 97
Security-as-a-Service (SaaS), 52
Security mechanisms
  access control, 180
  data computation, 179
  data storage, 179
  end devices, 179–180
  intrusion detection, 180
  network security, 178–179
Software-defined networking (SDN)
  big data in smart cities, 92–93
  implementation and analysis, 101–104
  proposed system architecture, SDIoT, 93–94
    application level, 100–101
    computational methods, 96
    data collection and intermediate levels, 95–96
    data management and processing level, 96–97
    HDFS system, 100
    message queue, 97–100
    named data network, 101
    threshold limit lalue, 97
  simulation results, 101–104
  Spark and GraphX, 101
Supply chain management
  adaptability technique, 152
  bolster exercises, 150
  cloud platform, IoT
    data quality, 164
    data security, 164–165
    investment justification, 166–167
    redesign process, 165–166
    technological change velocity, 165
  cost methodology, 152
  CSCMP, 148
  integrated value-added, 149–150
  inventory network, 148
  IoT applications
    big data and data analytics, 161–162
    cost reduction, 154
    data sharing and collaboration, 159–161
    innovation stage, 156
    logistics, 155
    monitoring and control, 158–159
    network architecture, 157
    open advancement, 157
    optimal distribution network, 154
    RFID labels, 153, 154
    streamline solutions, 162–163
    Things as a Service, 158
  leveraging effect, 151
  production network administration, 148
  quality procedure, 152
  SCOR model, 149
  time system, 152

T

Time division multiple access (TDMA), 37

W

Wireless medical sensor networks (WMSNs), 133