The Development of a Cloudlet Business Model Based on Stakeholders Service Guarantees

Mandisa Nomfundo Nxumalo
(201226712)
(B.Sc. Hons. Computer Science, University of Zululand)
University of Zululand
Department of Computer Science
Supervisor: Matthew O. Adigun
Co-supervisor: Ijeoma N. Mba

A dissertation submitted for the degree of

Master of Science' Degree in Computer Science, University of Zululand

2019

Declaration

This dissertation signifies the author's original work. This work is submitted by the author for the Master of Computer Science in the Department of Computer Science under the Faculty of Science and Agriculture at the University of Zululand. The author has not made any attempts to submit this work in the past for a degree or examination at any other university. All the materials used from existing works have been acknowledged and referenced.

Signature Date

Dedication

This work is dedicated to my late mother Thandekile Rachel Nxumalo.

Acknowledgements

I would like to first give praise unto the Lord my Almighty God for such wisdom, love and care. I would have not been able to make it this far if it was not by His grace. Thankful for the family He has given me for their care and patience with my journey. I am short of words to describe the respect and appreciation I have for Prof Matthew O. Adigun, Ms Ijeoma N. Mba and Mr Paul Tarwireyi. These individuals have played an important role in building my character and strength. Your guidance and support is much appreciated. I thank God for such people and in them I see God's love and care for me. It is true that family is not only by blood and I was able to experience that from the friends and colleagues in the Department of Computer Science. Their smiles, words of wisdom and care are the best one can ever ask for. To Prof. Matthew O. Adigun, Ms. Ijeoma N. Mba, Mr. Paul Tarwireyi, Mr. Isaiah Adebayo, Mr. Tosin Akinola, Ms. T. Ntuli, Mrs. D. Zibani, Mrs. N. Sibeko, Dr. Pragasen Mudali, Skhumbuzo Zwane, Sizo Ndlovu, Zibuyisile Magubane and others thank you. Iwould like to further extend my gratitude and acknowledgement to CSIR for their financial support throughout my Master's degree studies. Without their support my journey would have been uneasy. Thank you

Table of Content

List of Figures

List of Tables

A List of Acronyms

3G - 3rd Generation network

AP - Access Point

AR - Augmented Reality

CBA - Cost Benefit Analysis

CBM - Cloudlet Business Model

CBR - Cost Benefit Ratio

CSMA/CA - Carrier Sense Multiple Access with Collision Avoidance

DNS - Denial of Service

IDE - Integrated Development Environment

IoT - Internet of Things

ISP - Internet Service Provider

LAN - Local Area Network

LTE - Long Term Evolution network

MCC - Mobile Cloud Computing

NPV - Net Present Value

QoB - Quality of Business

QoE - Quality of Experience

QoS - Quality of Service

RAN - Radio Area Network

SLA - Service Level Agreement

SLIs - Service Level Indicators

SLOs - Service Level Objectives

SMEs - Small, Medium Enterprises

VMs - Virtual Machines

WAN - Wide Area Network

Wi-Fi - Wireless Fidelity

WLAN - Wireless Local Area Network

List of Publications

Nxumalo, M.N., Adigun, M.O. and Mba, I., 2018, August. An envisaged SLA- based Cloudlet Business Model for ensuring Service Guarantees. In 2018 Inter- national Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD) (pp. 1-4). IEEE.

Nxumalo, M.N., Mba, I.N. and Adigun, M.O., 2018, December. The Value Perspective to Deploying a Cloudlet Business Model for SMEs. In 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC) (pp. 1-6). IEEE.

Nxumalo, M.N., Adigun, M.O. and Mba, I.N., 2020. The Impact of Service-Level Agreement (SLA) on a Cloudlet Deployed in a Coffee Shop Scenario. In Fourth International Congress on Information and Communication Technology (pp. 397-406). Springer, Singapore.

Abstract

In this study, a Service Level Agreement (SLA)-based Cloudlet business model was developed to facilitate the easy deployment of a Cloudlet instance by Small Medium Enterprises (SMEs) that ensures service guarantees by the edge. The development of the model is due to the lack of Cloudlet Business Models to assist and guarantee SMEs that the deployment of a Cloudlet can yield desired profits.

The model consists of three role-players or stakeholders namely, the Cloudlet consumer, Cloudlet owners and Internet Service Providers (ISP) or Cloud Providers. The roles and interactions between the role-players are captured in an SLA. The SLA facilitates the provisioning and management of Cloudlet resources to ensure service guarantees and avoid deployment failures. The management and provision of Cloudlet resources can result in high throughput in the network, improving consumer Quality of Experience (QoE). The SLA defines the service description, service objectives, cost, and penalties and is imposed on consumers using a disclaimer page.

The feasibility study of the developed model based on a coffee shop scenario was conducted using a Cost-Benefit Analysis (CBA) tool. The findings show that a coffee shop can gain about 164 cents per rand spent on Cloudlet operational costs over a period of 3 years. This can surely help SMEs gain financial stability and profit. Also, to demonstrate the effectiveness of an SLA on the Cloudlet Business Model, a coffee shop-based SLA simulation was conducted using CloudSim Plus tool. The results indicated that the provisioning of efficient resources in the network had a direct impact on the success of the SLA. The SLA's success had an impact on both service guarantees and operational cost minimisation. Therefore it is concluded that an SLA-based Cloudlet business model proposed in this research ensures that stakeholders' service guarantees are delivered.

Chapter 1

Introduction

1.1 Overview

A Cloudlet being a single-hop, resource-rich computer or cluster of computers, enables mobile users to access cloud resources using cheaper Wi-Fi technology or mobile wireless Local Area Network (LAN). (Pang et al., 2015, Gao et al., 2015). The Cloudlet was introduced by Mobile Cloud Computing (MCC) to reduce latency encountered by response sensitive applications such as online gaming and video conferencing. Using these merits, this research pays close attention to the deployment of a Cloudlet technology as a means of providing value and benefits for all parties involved in the Small Medium Enterprises (SMEs) environment. The SMEs under discussion are not limited to the Internet cafés, coffee shops, malls, libraries, and doctors' offices, SMEs are described as enterprises that have a high impact (such as contributing to job creation and poverty alleviation) on the economic growth of the country.

SMEs encounter high failure rates due to the lack of funding and access to finance, which impacts their ability to grow and operate (Cant, and Wiid, 2013, Ewe, 2017). To gain financial stability an SME can deploy a Cloudlet as a means of attracting more customers, improving business marketing strategies, business-exposure, core product purchase patterns and profit generation (Shaukat et al., 2016, Pang et al., 2015). The above is feasible because deploying a Cloudlet is inexpensive and maintenance free (Pang et al., 2015). Amongst other merits possessed by a Cloudlet is computation offloading which is done through enabling mobile devices to offload computation using cheaper Wi-Fi technologies at a one-hop, high bandwidth and low latency connection (Verbelen et al., 2012, Panigrahi et al., 2015).

The sequence of events in computation offloading occurs in the sense that the mobile devices connect to the Cloudlet in their proximity and begin to initiate service requests (storage, CPU, bandwidth, and capacity) to the Cloudlet (Pang et al., 2015). The Cloudlet then processes the requests and sends back the results to the mobile device.

The fact that using a Cloudlet to access cloud services reduces latency and increases throughput has been recognised by both mobile users and businesses (Jararweh et al., 2014, Pang et al., 2015). The supposed benefits derivable from a Cloudlet can be further leveraged by turning a Cloudlet into a smart data sensor of small data (Gordon, 2014, Lindstrom, 2016). Small data refers to the information that defines and articulates an individual user preferences based on their internet browsing history. This means that the individualized small data can be extracted with the owners' permission as metadata information from a Cloudlet cache; in contrast to big data, where extracted metadata is used for the benefit of the public (Krueger , and Dix, 2014). A smart Cloudlet is described as one that is context-enabled, fulfils expected stakeholders' service guarantees while increasing the Quality of Experience (QoE) of consumers. QoE in this study refers to the measure of amusement or frustration of consumer experience with the Cloudlet. In this study, context refers to the repository of operational data generated by the consumer device before, during and after the interaction between the mobile device and the Cloudlet called a user browser history.

A context-enabled Cloudlet refers to a Cloudlet that uses mobile browser history or small data to provide customised online services. The consumer device, being the source of the small data sensed by the Cloudlet, is restricted to relevant input possibilities and relevant data (Loke, 2012, Kotevska et al., 2016). Studying a context-enabled (detect, capture, analyse and present to the user) Cloudlet is a pre-requisite to determining what services can be suggested to the consumer and at what price (Loke, 2012, Mor, 2014). Therefore, a context-enabled Cloudlet is expected to enhance the derivation of an attractive Cloudlet Business Model (CBM).

The CBM is developed based on the Cloudlet as a Service Model and is aimed to facilitate the deployment of Cloudlet technology by SMEs. A CBM consists of three role-players, namely the Cloudlet consumer, Cloudlet owner (SME owner), and Internet Service Provider or Cloud Provider. These role-players are bound by a contract called a Service Level Agreement (SLA). The SLA envisions data sources are mobile device users that are willing to exchange their browser history data for Cloud resources residing on SMEs-deployed Cloudlet instance. Apart from the Cloudlet consumers' role, all other participants (Cloudlet owner, Internet Service Provider or Cloud Provider) in the CBM impact the operations and the Quality of Service (QoS) offered which in return determines the Quality of Business (QoB). QoS refers to the level of service a provider is willing to offer to consumers and is specified in the SLA. On the other hand, QoB refers to the service feedback by consumers, which is measured by the rate of consumer interest to stay on SMEs or consume the offered services. The lower the QoS the poorer the QoB. The SLA specifies the service parameters (Service Level Objectives), Quality of Service (expressed as Service Level Indicators) and related issues such as penalties and benefits that can be encountered during the negotiation process (Buyya, and Bubendorfer, 2009, Xiong, and Chen, 2015).

In this study, it is assumed that the SLA parameters and constants (SLOs and SLIs) are limited to those defined as elements in a price function; and their outcomes during the execution process of a service is mapped using a utility value. The richer the QoE of a consumer, the higher the utility value derived. During negotiations of an SLA, the consumers' perspective is that a service provider should offer better QoS for a low price; but a service providers' perspective is that, if the consumer wants higher QoS, they should be willing to pay more (Buyya, and Bubendorfer, 2009). To achieve fairness in the SLA negotiation, outcome-based pricing is selected on the basis that the price function in the SLA will be written only in terms that are visible to both parties and not some randomly created values by the service provider (Buyya, and Bubendorfer, 2009, Xiong, and Chen, 2015). This will ensure that the outcome satisfies all role-players' expectations and eliminates payment for encountered SLOs violations or breaches of service guarantees.

1.2 Problem statement

The Cloudlet paradigm subscribes to a number of Mobile Cloud Computing business advantages such as extending battery life, data storage capacity, and processing power as well as improving reliability, availability, and dynamic provisioning (Duro et al., 2015, Mach, and Becvar, 2017). However, the commercial deployment (Satyanarayanan et al., 2015) of a Cloudlet remains a challenge. This is because very little is known about Cloudlet Business Models. It can be considered a risk for SMEs such as an Internet café to trust that the deployment of a Cloudlet on their premises can lead to maximising profit and ensuring service guarantees for their consumers. The question remains: what sort of service guarantees should stakeholders or role-players expect from one another when deploying a Cloudlet, if deployment was to optimise profit, attract consumers (by a prospect of higher quality of experience) and enforce compliance to the Service Level Objectives (SLO) stated in the Service Level Agreement (SLA)? This work, therefore, focuses on ensuring that both SLA-based discovery and negotiation processes of a Cloudlet business model are matched to an efficient price function that fulfils service guarantees for consumers.

1.3 Motivation of the study

Cloudlet technology envisages the decentralised ownership by SMEs such as coffee shops or doctors' offices where they could serve as small data sensor sources. The technology is premised on simplified or fully autonomous management, which makes it a strong business proposition (Satyanarayanan et al., 2009). Self-management corresponds to the appliance model of computing, making its deployment easy and beneficial in SME premises (Kotevska et al., 2016, Satyanarayanan et al., 2009). However, the Cloudlet business model is one aspect of the technology that the research community has neglected until now. Unresolved issues include but are not limited to: deciding whether the business model should be bottom-up (benefits a consumer) or top-down (benefits the provider), the lack of existing Cloudlet Business Models to support its deployment, and finally the use of a price function that does not align with the resource specification in the SLA document (Satyanarayanan et al., 2009). Therefore, this study is expected to yield outcomes that will help motivate SMEs to explore the bottom-up deployment of the Cloudlet Business Model.

1.4 Research questions

The research questions that this work addressed are:

a. How will the SLA-based Cloudlet Business Model benefit from Cloudlet as a small- data-sensor?

b. How can the SLA's (SLO and SLI) of role-players in a Cloudlet Business Model be captured in the price function (s)?

c. Which price function can be recommended for an SLA-based Cloudlet Business Model?

1.5 Research Goal and Objectives

1.5.1 Goal

The goal of this study is to develop an SLA-based Cloudlet business model that uses an appropriate price function to negotiate the satisfaction of all role-players in the Business Model.

1.5.2 Research Objectives

a. To investigate existing SLA-based Cloud business models or similar models that can be aligned to a Cloudlet environment and price functions that fit such an environment.

b. To explore how the use of small data sensors can be beneficial to an SLA-based Cloudlet business model.

c. To design a proposed Cloudlet business model that integrates SLAs of role-players with appropriate pricing functions.

d. To evaluate the effect an SLA can impose on a CBM using a coffee shop scenario.

1.6 Research Methodology

The research methods given below provide an overview of how the research objectives of this study were achieved:

a. Literature-review

Existing studies from different scholars regarding Cloud, Cloudlet and factors involved in deploying a Cloudlet or any edge computing technology such as business models, performance benchmarks, and the importance of concepts: QoE (Quality of Experience) and QoS (Quality of Service) for provisioning a better SLA were investigated in order to build a state-of-the-art for this study.

b. Simulation

This methodology used the information and knowledge gained from the literature to conduct a simulation of an SLA to prove its functionality in ensuring service guarantees and resource management in the CBM. The simulation was conducted based on a coffee shop scenario.

The simulation involved metrics such as the *task completion time* in milliseconds to process a batch of 150 tasks against an increasing number of virtual machines (10-50 VMs) which impacted the amount of *CPU utilization*, and task *wait time* which further influenced the total network *throughput* that correlates with the obtained Service Level Objectives (SLOs) percentage per simulation session. A SLO percentage lower than 100 % signified a Service Level Agreement (SLA) violation, and using the *expected task* *completion*, user session was increased to ensure better QoE.

1.7 Contribution to the Body of Knowledge

The contribution of this study to the existing body of knowledge is two-fold: First, while existing Cloudlet related business model studies have focused on solving the resource-poverty aspect of the Mobile device, this study is paying first-of-its-kind attention to how device owners could benefit from a business model-assisted trade-off of their own small data. Second, the trade-off is achieved based on an efficient pricing function that ensures service guarantees offered by stakeholders are enforced.

1.8 Dissertation Synopsis

The rest of the dissertation document is organised as follows:

In Chapter 2, a background study of edge computing technologies is discussed with focus on the Cloudlet technology. The Cloudlet background covers Cloudlet features, Cloudlet as the Service Model and the use of SLA to achieve Cloudlet as a Service Model.

Chapter 3 provides related work to this study pertaining to SLA-Based business models used in Cloud and Grid computing, the applicability of context-awareness in a Cloudlet, types of competition-based price functions used in the Cloudlet environment and the simulations done based on the SLA.

In Chapter 4, a discussion that describes the development of a CBM is covered. The discussion covers the role-players, model components, processes, use-case scenario and the value proposition of a CBM.

Chapter 5 provides evidence of the feasibility of the Cloudlet Business Model using the coffee shop scenario.

Chapter 6 presents the discussion of the results obtained from simulating an SLA as task requests are initiated and processed in the network using several performance metrics. The simulation is designed based on the coffee shop scenario.

Chapter 7 provides the conclusion, limitations, and future work

Chapter 2

Background

2.1 Introduction

The increase in the intelligence and usage of mobile devices to access Cloud resources has led to an increase in problems encountered by mobile users such as limited storage and computing power (Yu et al., 2018). To overcome the above-mentioned challenges a new field emerged out of the integration of Cloud Computing with Mobile Computing called Mobile Cloud Computing (MCC) (Pang et al., 2015, Mach, and Becvar, 2017). The MCC conceptually moves the computation power and storage from a mobile device to a remote controller so these resources can be remotely accessed (Raei, and Yazdani, 2017). This means that the data storage and computation happen inside the remote Cloud and the results are returned to the mobile device. In MCC, the mobile devices utilise a Cloud service via either a Wi-Fi or cellular operator service. However, the process of utilising the Cloud service results in WAN (Wide Area Network) latency which triggers more delays due to the distant location of the Cloud from mobile users. Latency is less favourable for time-sensitive applications such as online gaming, video conferencing, image processing and others. To overcome latency related issues, edge cloud servers such as Fog, Mobile Edge and Cloudlet computing were introduced by MCC (Pang et al., 2015). These edge servers share a common aim that is to bring the Cloud services to the proximity of mobile users. The edge paradigms also play a vital role in the interaction and transmission of real-time data between devices such as sensors in the Internet of Things (IoT) network (Ai et al., 2018). For instances in a case where a chemical vat safety mode is breached, informative action(s) must be taken with immediate effect to avoid catastrophic events. However, this might be delayed by the time (trip) it takes for temperature readings to pass from the edge devices (sensors) to the Cloud for processing and might limit chances of avoiding catastrophic events. The usage of edge paradigms which help analyse data closer to where it is generated decreases latency issues, furthermore, giving time to generate and implement ideas that can avoid high damage occurrence for the above-stated example. Although the edge paradigms share a common vision, they possess different sets of features that set them apart. The below discussion covers the description of the different features possessed by edge paradigms, use-cases, and their open issues.

2.2 Types of Edge Computing paradigms

There are three types of edge computing paradigms, namely fog computing, mobile edge computing, and the Cloudlet. Below captures the discussion of the above-stated paradigms covering their features, deployment models, differences and challenges. (Computing, 2015, Dolui, and Datta, 2017, Bonomi et al., 2012, Borcoci, 2016, Pang et al., 2015, Ai et al., 2018, Dou, and Heinzelman, 2013, Yu et al., 2018)

Fog Computing is defined as a highly virtualised three-tiered architecture a renders the high amount of computing, storage, and management of networking services amongst end devices and Cloud data centres. The multi-tiered architecture consists of a Cloud instance, fog nodes (edge servers, gateways, switches or routers) and edge devices (smart sensors or mobile devices). It ensures high efficiency by reducing the amount of data that must be transported to the Cloud for processing, analysis, and storage. It is mostly used in IoT applications due to the ability to overcome Cloud limitations such as low latency, mobility, and location awareness. The applications of fog computing cover a number of critical IoT services, and applications such as smart cities, health care, smart grids, and connected cars.

Mobile Edge Computing (MEC) was introduced for the purpose of addressing MCC challenges such as high latency, security, low coverage and lagged data transmission through deploying Cloud resources to the edge within a Radio Access Network (RAN). MEC brings about several advantages to mobile users such as low latency, energy savings, geographical awareness, context awareness, and strengthened security and privacy-related issues. The MEC servers or nodes are located within the Radio Network Controller or a macro base station, running multiple instances of MEC hosts. The MEC hosts are managed by a Mobile Edge Orchestrator that manages information on the services offered by each host, available resources and network topology while also handling the Mobile Edge applications. MEC is used to accelerate content, services, and application responsiveness by the edge. The applications of MEC involve video analytics, distribute content and denial of service (DNS) caching, augmented reality (AR) service delivery, and application-aware cell performance optimisation. It highly promotes mobility and can be deployed on LTE macro base stations, multi-technology (3G/ LTE) cell aggregation sites, and at 3G Radio Network Controller sites.

Lastly, a Cloudlet is defined as a cluster of computers that are well connected to the Internet, well equipped with resources available for use by proximate mobile devices. A Cloudlet is found in the middle of a three-tiered architecture (Cloud-Cloudlet-mobile devices) acting as a data centre in a box running virtual machine instances that are capable of providing resources to mobile devices in real-time over a Wireless Local Area

Network (WLAN). The Cloudlet architecture provides one-hop access with low latency and high bandwidth service allocation for applications. Furthermore, it can be installed indoors (doctor's office, Internet cafe, community mall, coffee shop, and others) or outdoors (base stations). It is used to support use cases such as in the health sector, security sector, and consumer services discretionary to cognitive assistance.

Fig. 2.1, shows different deployment scenarios of edge paradigms that take care of a video surveillance use case, which uses 22 Mbps of Internet connection.
The assumption is that cameras, push video frames to the Cloud for analysis whenever motion is detected. Depending on the configured triggers, if the frame is interesting a notification is sent to the user. In cases where there is a network outage, the connection is terminated and no notification is sent back to the user.

In all deployment scenarios, the network diagram layout is similar. However, the processing of video frames is done differently as follows: (Bonomi et al., 2014)

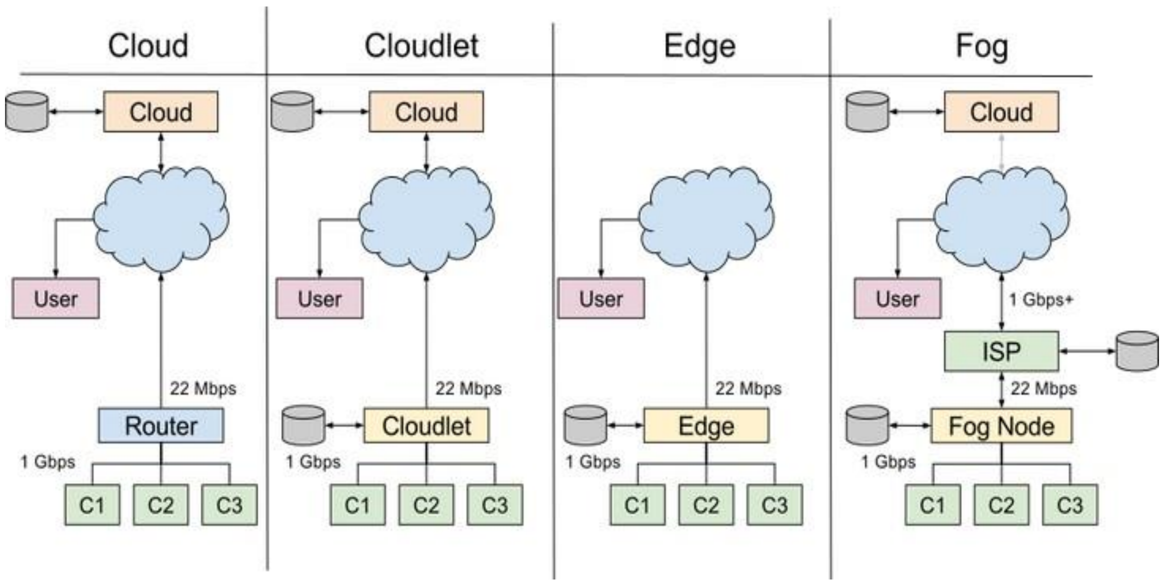

Figure 2.1: Edge Computing Paradigms (Bonomi et al. (2014))

For Fog, Cloudlet and Mobile Edge, the video frames are pushed to the edge servers for further instructions. In the Cloudlet use case the data frames are divided into sub-frames and pushed to the Cloud for further analysis and the notification is sent back to the user. In Fog and Mobile Edge, the processing of the video frames is taken care of at the edge through using computer vision functionality similar to Cloud which notifies the user. However, in Fog the processing and notification generation can occur within the edge because it has additional storage. In cases where there is a network outage, the connection is not terminated, but the notification is not sent unless there is an installed backup connection like a cellular network.

Therefore, using edge servers for processing data is more efficient than directly using Cloud servers due to latency. A summary of the distinct variation between the edge paradigms is shown in Table 2.1. The Table covers a summary of characteristics such as types of edge nodes used, where the nodes are located, the proximity of these nodes, architecture type, availability of data analytics and software architectures involved (Bonomi et al., 2014, Yu et al., 2018, Pang et al., 2015, Raei, and Yazdani, 2017). Through literature review it was obtained that the most common open topics amongst the edge paradigms are possible operational/ deployment models, determining the requirements for security and privacy issues. However, this research has a close interest in the deployment models SMEs can adopt for deploying a Cloudlet edge server.

2.3 Cloudlet Technology

All the edge servers bring about low latency to the edge, but the edge server of interest to this study is a Cloudlet (shown in Fig. 2.2), due to the fact that SMEs can deploy a Cloudlet to stabilise finances while enabling consumers to enjoy a low latency, high bandwidth edge server for improving the process cycles of time-sensitive applications. This Section covers both the Cloudlet advantages and service models found in the literature (Panigrahi et al., 2015, Satyanarayanan, 2015, Bilal et al., 2018, Hou , 2018, Shaukat et al. , 2016)

The Cloudlet architecture enables the reduction of latency in response-sensitive applications (such as augmented reality, online gaming, face recognition, and crowdsourcing video) processing by using a one-hop network connection.

The features of a Cloudlet bring about the following advantages:

1. **Easy Deployment:** The stateless characteristic of a Cloudlet simplifies their management for SMEs. A Cloudlet is self-managed requiring little or no professional administration or operation. This means that the adding or replacement of a Cloudlet requires a moderate setup or configuration effort.

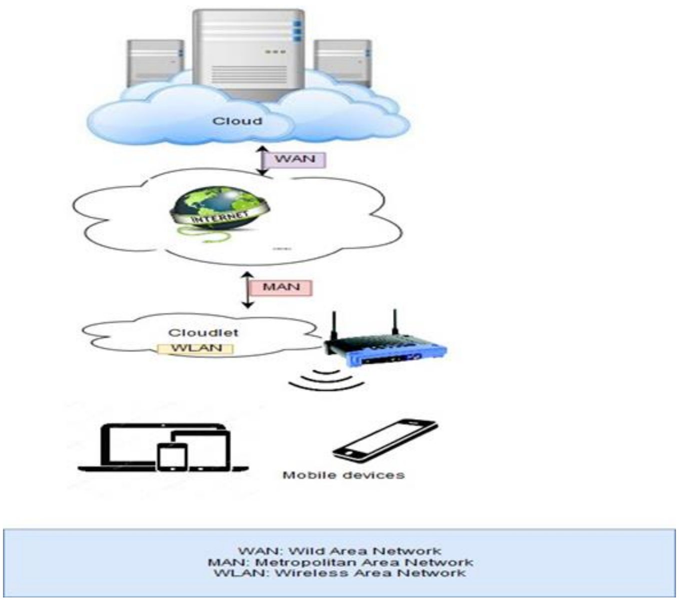

Figure 2.2: Cloudlet Architecture

Table 2.1: Comparison between Edge Computing Paradigms

Feature	Fog Computing	Mobile Edge Computing	Cloudlet
Edge nodes	Routers, Switches, Access Points and Gateways	Servers operating in base stations	Data centre in a box.
N-tier hierarchy	N=3	N=2 or 3	N=3
Online data Analytics	available	Not available	Not available
Location of edge nodes	Between end devices and Cloud	RNC/ Macro Base station	Within a business or other user-focused locations
Proximity	One or multiple hops	One hop	One hop
Access Mechanism	Bluetooth, Wi-Fi and Mobile Networks	Mobile networks (LTE, 3G, others)	Wi-Fi
Support for Context awareness of applications	Medium	high	Low
Software architecture	Fog abstraction layer based	Mobile orchestrator-based	Cloudlet agent-based

2. **Security improvement:** The proximity of a Cloudlet by the edge makes it more resilient against Denial of Service (DoS) attacks. Also, since the Cloudlet owner can restrict the range of end-to-end communication, it prevents distance attackers from accessing network traffic information.

3. **Resilience:** Regardless of the weak connection that may occur between the Cloudlet and the distance Cloud, the Cloudlet can still offer reliable Cloud computing services to end users.

4. **Cost-effective:** Since a Cloudlet stores soft state data copies in its cache, the occurrence of a Cloudlet failure in the network has not resulted in incurable

events.

5. **Offline availability:** The Cloudlet allows mobile devices to offload computation to it without the need for an active Internet connection. This means that Cloudlet-offered services can be accessed anytime.
6. **Alleviation of Cloud load:** The Cloudlet reduces the load in the Cloud by processing some of the Cloud applications such as time-sensitive applications.
7. **High bandwidth:** The Cloudlet uses a Wireless Local Area Network (WLAN) to assist mobile devices to gain access to Cloud services. The bandwidth of WLAN technologies ranges from 2 Mbps - 400 Mbps (Modh, and Patel, 2013, Kbar, and Mansoor, 2006). One of the most popular technologies for WLAN is Wireless Fidelity (Wi-Fi) which is obtained through connecting to devices such as Wi-Fi-enabled Access Points (AP) or routers. The AP is in the form of a Cloudlet-Wi-Fi and the mobile devices use 802.11n to connect to the configured private network.

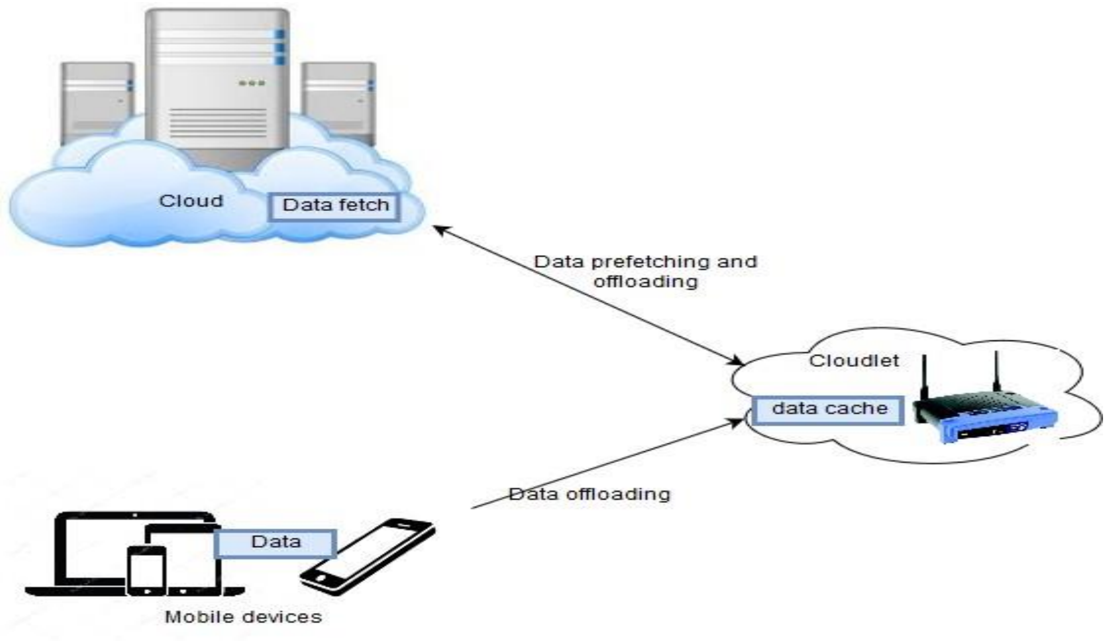

Figure 2.3: Data Cache and prefetching for ensuring customised services by the edge

Through using WLAN mobile devices are able to efficiently and effectively access Cloud services.

8. **Efficient data access:** Efficient data access is achieved by allowing a Cloudlet to cache mobile device data in order to provide customised online services. This means that a Cloudlet is context-aware of consumers' online interests. Data caching and pre-fetching are two terms that come into play in achieving efficient data access (Tang et al., 2008, Dar et al., 1996) and these terms are discussed on the following section.

2.3.1 Data caching and pre-fetching

The term data caching and pre-fetching are important concepts in ensuring efficient data access on a Cloudlet. Data caching refers to the ability of a Cloudlet to capture jobs requested by end users on its storage for later use. While data pre-fetching refers to the ability of a Cloud instance to download and offload services of preference to end users based on the content cached on the Cloudlet.

The pre-fetch component is influenced by the data cached on a Cloudlet. Data caching resource limitations. A data cache component in the Cloudlet stores the offloaded data to ensure future alleviation of data access without encountering high connectivity delays. Since the Cloud runs virtual applications of mobile devices, a data-fetch component on the Cloud downloads end-user-based relevant services and offloads it back to the Cloudlet. Lastly, the Cloudlet then offloads part of the data to the mobile device. The caching and pre-fetching mechanisms improved service delivery on the network, thus improving the end users QoE. The caching of data on the Cloudlet requires management mechanisms to avoid storage limitations. For example, an architecture to reduce latency and bandwidth utilisation when accessing data from a Cloud service was introduced (Zhijun et al., 2017). The framework made use of a data replacement strategy in a Cloudlet cache to pinpoint and replace chunks of data that have not been accessed within a certain period of time. The function of a data replacement technique is to determine if the requested data from mobile devices is stored in the Cloudlet cache. If the requested data is available in the Cloudlet cache, it increases the unit access of the data on the cache. Depending on the set threshold of data unit access, the data might be removed or stored in the cache to avoid storage of useless data. The types of data replacement algorithms include cost-based, locality-based, semantic-based and utility-based data placement algorithms that can be implemented on a Cloudlet interface (Podlipnig, and Böszörmeny, 2003). Such strategies minimised the cost of adding storage and maximised efficiency on the network.

2.3.2 The use of Wireless Local Area Network (WLAN) by a Cloudlet

A Cloudlet makes use of a Wireless Local Area Network (WLAN) also known as a Local Area Wireless Network (LAWN), a type of connection that allows mobile users to gain access to network resources without the need for Internet wires or cables. The connection is established through a Local Area Network (LAN) that uses radio waves to transmit wireless signals to mobile devices. The WLAN follows the IEEE 802.11 Wi-Fi standards that defines the WLAN specifications between two endpoints (wireless client and base station or between two wireless clients). Furthermore, WLAN uses Ethernet protocols and Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) to enable efficient path sharing (Hiertz et al., 2010). The advantages of WLAN are not limited to the list below: (Ala-Laurila et al., 2001, Wu et al., 2002)

- **Cost**: Due to the elimination of cables, the operation of WLAN reduces the cost required for resource purchases. Also, if compared to 3G networks the WLAN such as Wi-Fi cost less in terms of bandwidth.

- **Data privacy**: The availability of and accessibility to WLAN depends on the proximity of the mobile device. Only connected devices in the proximity of the WLAN can gain access to its data. Due to the improvements and advancements of WLAN, Wi-Fi technology is able to provide robust security.

- **Scalability**: it allows the expansion of the network by using existing equipment, where-as wired networks might need the addition of Internet cables or other equipment to extend the WLAN signal transmission.

- **Mobility**: Mobile devices can be connected to the WLAN as long as they are in its proximity.

- **Easy setup**: It is easy to configure a network that uses WLAN due to the fact that it does not require the installation of multiple cables per device located on the network.

The merits possessed by WLAN makes it a highly demanded technology in various environments such as in companies and home setups. One of the most popular technologies for WLAN mentioned above is Wireless Fidelity (Wi-Fi) which is obtained through connecting to devices such as Wi-Fi-enabled Access Points (AP) or routers. The AP or router can be in the form of a Cloud-Wi-Fi and Cloudlet-Wi-Fi. In both Cloud-Wi-Fi and Cloudlet-Wi-Fi, the mobile devices use 802.11n to connect to the configured private Wi-Fi AP, which is connected to the enterprise Ethernet (Kbar, and Mansoor, 2006, Modh, and Patel, 2013). However, the difference remains on the routing of network traffic. In Cloud-Wi-Fi the network traffic goes via the enterprise Ethernet to an Amazon AWS site. On the Cloudlet-Wi-Fi the network traffic goes to the Cloudlet that is on the same Ethernet proximity as the Wi-Fi AP.

A comparative study between the Cloud-Wi-Fi and Cloudlet-Wi-Fi was conducted to evaluate which setup could improve response time and energy consumption on the network (Hu et al., 2016). The setup of the comparison for performance evaluation took into consideration response-sensitive applications namely, FACE (face recognition), MAR (augmented reality) and FLUID (physics-based computer graphics). The results prove that offloading computation to the Cloud-Wi-Fi or distance Cloud reduces network performance and increases energy costs if compared to Cloudlet-Wi-Fi results.

The routing of network traffic to the Cloud distance Cloud instance increases delays in application response time. The delays in application response time are due to issues such as network latency, data transmission time, processing time at the data centre and processing time on the mobile device. Cloudlet-Wi-Fi is the advancement of Cloud-Wi-Fi WLAN technology, it reduces the above by processing at the edge using a one hop connection to the mobile devices (Modh, and Patel, 2013).

2.4 Cloudlet as a Service Model

This study focuses on the Cloudlet as a Service model. This is because a Cloudlet deployed in an SME will act as a value-added service that provides more incentive to the SME customers to increase their visit and stay. According to this model, the Cloudlet augments a mobile device by offering services such as data processing, caching and partitioning or scheduling of jobs (Shaukat et al., 2016).

Due to the limited resources on a mobile device, a Cloudlet enables mobile devices to offload computation for processing. This helps the mobile device to save resources such as battery-life when processing a large amount of data. The data processing occurs with minimal delays when compared to the distance Cloud. When data has been processed, a Cloudlet utilises its internal storage to cache and store data for later use in cases when the same data request is sent for processing.

The Cloudlet cache the data to provide services of preference to end-users. This is achieved by the use of intelligent prediction and behaviour models that use historical data to predict and provide end-user queries. Lastly, a Cloudlet providing a caching service does not require a user application host on it. After the data has been cached, this means that the data is available for external use such as analyses. The Cloudlet uses scheduling algorithms such as queuing models to delegate tasks across multiple Virtual Machines or servers in a network. This is done with an aim to increase the overall throughput on a network, and reduce latency related to task processing. Lastly, the partitioning of the requested tasks requires the internal knowledge of its metadata.

2.5 SLA for Cloudlet as a Service Model

Ensuring the effective provision of data processing, caching and partitioning/ Scheduling of jobs as a service in a Cloudlet requires the use of a management policy. The SLA as shown in Fig. 2.4 is a service management policy issued by the service provider to the consumer. It defines and regulates the relationship between the service provider and service consumer. The said, definition and regulations is achieved through: (del R´ıo-Ortega et al., 2015, Patel et al., 2009, Xiong, and Chen, 2015, Xiong, and Chen, 2015) The identification of the role-players or parties (service provider and consumer); period of the contract (start and duration); identification and description of the Service for provision to consumers; Service Level Indicators (SLI): How service provisioning will be monitored; expected quantity or level of service to be provided by service providers to consumers; countermeasures in place in case of service violations; lastly, the price of service provisioning (price function/ model) and proof of agreement for all involved parties on the SLA negotiations.

The price functions for charging for the provisioning of services are categorized into dynamic and fixed price functions that can either be in a state of monetary or non-monetary value exchange functions (Ibrahimi, 2017). The dynamic price function is mostly influenced by the market value and the response of consumers towards product consumption which for example involves supply and demand. Opposite to the dynamic price function, the fixed priced function uses pay-per-use or subscription pricing. This means that consumers pay based on the amount of service consumed.

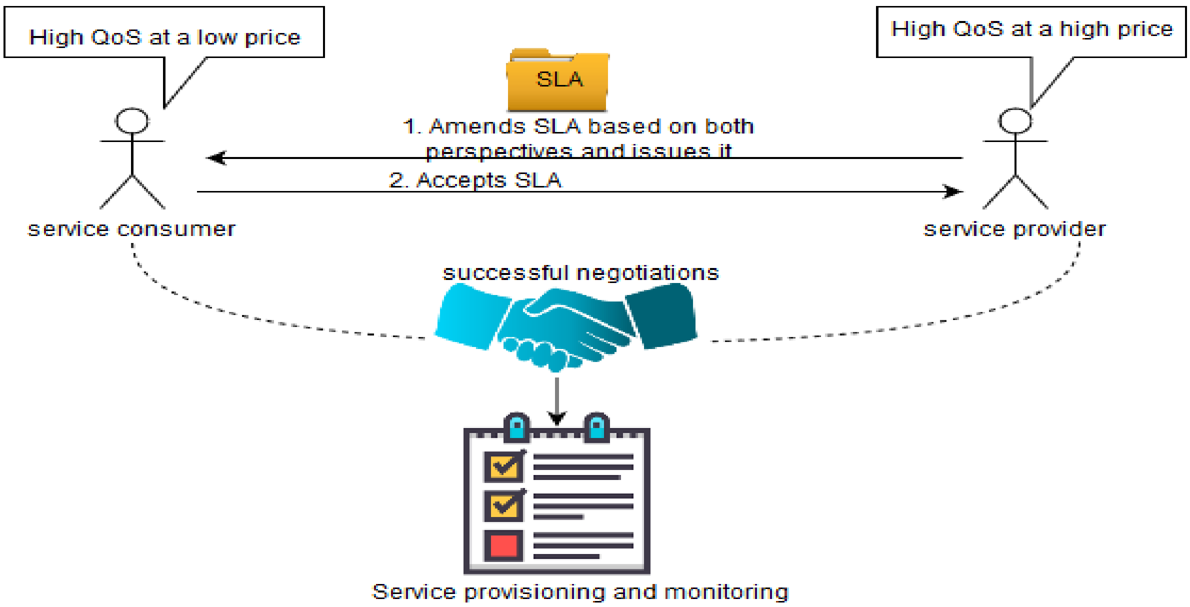

Figure 2.4: Service Level Agreement (SLA) Negotiation

An effective SLA is unambiguous (all terms as mentioned above are clearly presented at the SLA) and takes into consideration the perspectives of the involved parties to ensure fairness and service guarantees. A perspective of a service consumer is that the service provider should provide a service of high quality at a low price. On the other part, a service provider's perspective is that for consumers to receive services of high quality they must be willing to spend more. The reaching of an SLA agreement occurs when both parties understand and are able to adjust the SLA to ensure fairness to all. The use of an SLA ranges from ensuring efficient service provisioning on the network to the management of resources in order to improve consumer QoE.

2.6 Chapter Summary

This Chapter started by introducing the edge computing paradigms namely fog, mobile edge and Cloudlet computing technologies. Then outlining the difference between the edge computing paradigms using the processing of video frames example. The discussion was then narrowed down with focus on a Cloudlet technology. The Cloudlet background covered the characteristics of a Cloudlet, how a Cloudlet is used as a Service model and how SLA can be used to achieve a better service model with regards to a Cloudlet technology. The following chapter focus on the work that has been done by other authors with respect to context data, service model and the benefit of using SLA in a context-aware Cloudlet service model.

Chapter 3
Literature Review

3.1 Introduction

The previous Chapter provided a descriptive background on the important key points of this research such as data caching for ensuring context awareness, Cloudlet as a service model and the usage of SLAs to improve the provision of services. This chapter extends the above by discussing the literature around the key points. Due to the lack of Cloudlet business models, this chapter looks at the existing grid or cloud business model deployment mechanisms to build a background for developing a Cloudlet business model.

The literature covered in this chapter includes the review of strategies to enable context-awareness on a Cloudlet, cost functions used in a Cloudlet, SLA-based business models, and SLA-based simulations conducted in a network environment.

3.2 Enabling Context-Awareness in a Cloudlet technology

The enhancement of a Cloudlet's performance in order to minimise the cost of its resources and efficiently save the resources requires that a Cloudlet be context-aware. This chapter reviews different works which elaborated on the use of context-systems on a Cloudlet and on the Cloud, with emphasis on how these systems can help the efficient functioning of the network.

Many data-sensing applications and technologies that deploy context-awareness have been developed in different disciplines such as health care services, government, industrial and other sectors to help collect information from the surroundings (data sensors) that will be beneficial for those who need it. An example that can be derived from this includes that of augmented reality and face recognition technologies, whereby an image from a mobile device is transmitted to the Cloudlet over the Internet using Wi-Fi (Botta et al., 2016, Satyanarayanan , 2015). The transmitted image is then processed at the Cloudlet using a computer vision algorithm. Features contained in the image corresponding to their identity are sent back to the mobile device. In a face recognition request, the identities of the person in the image are returned to the mobile device (Botta et al., 2016).

The sensing of data can be obtained through the usage of big data or small data concepts. The most popular in Cloud being the big data concept. In big data, to track the state of an entity involves the integration of multiple sensors communicating with each other using a multi-hop communication to provide a joint result (Botta et al., 2016, Neves et al., 2016). The collection and processing of sensed data from mobile devices is done by using data science techniques and technologies (Antonić et al., 2016). The data is then analysed and distributed using Hadoop (Alam, and Shakil, 2016). The distributed and analysed data by Hadoop (an analytic software) is stored on the Cloud to enable mobile users to access it. Big data has many advantages however, the concept still suffers from challenges such as resource efficiency (power considerations, storage capacity, processing capabilities, and bandwidth availability) (Neves et al., 2016).

Arising from the above-mentioned limitations of big data, this study focuses on using a small data approach which does not require highly complex systems for data-intensive analysis, large storage or bandwidth for its access and usage. The concept of small data refers to the sensed information obtained from a mobile device, detailing the user context and preferences. The choice of small data rather than big data aligns with the limitation of storage, computing capabilities possessed by the Cloudlet (Rawadi et al., 2014) and the mere fact that Cloudlet is capable of generating analytics for the small data.

The question still remains: how will sensed data be collected and analysed by a Cloudlet? In the work done by (Loke, 2012), they explored the usage and deployment of a sensor-Cloudlet and context-Cloudlet by means of providing sensor and context information on an on-demand basis. These concepts essentially involve the composition of context-aware systems such as an abstraction over sensors, context interpreters and situation reason engines with applications. This was done to provide a starting point for operators that might be used in composing resources for Cloudlets.

The study by Loke, 2012 did not address issues such as the discovery of resources to fulfil user's needs, which can improve the QoE.

The mobile crowd sensing technique is a high contributor in the IoT (Internet of Things) applications (Guo et al., 2015, Liu et al., 2018). The mobile crowd sensing technique is formed by a wide variety of sensing models, in which each sensing model collectively shares data and extracts information to measure and map the occurrence of a common interest (Guo et al., 2015, Liu et al., 2018). The application of this technique is commonly deployed on a mobile or a personal device that can be used to provide sensor data about the physical environment to the mobile application server. It requires multiple participants in sensing the surrounding environment using sensing devices with built-in sensors. The continuous increase in the amount of data generated by the sensing devices leads to a heavy traffic load and bandwidth constraints in a network. A survey conducted by (Liu et al., 2018), explores how existing work on mobile crowd sensing strategies are used to reduce resource cost and improve QoS. The above is achieved through the adoption of a deduplication method used to eliminate redundancy in the sensed data.

A Real-time framework for Cloudlet-based architectures was introduced in (Kotevska et al., 2016). The framework is based on prediction models that provide analytical methods such as integration, aggregation, processing of real-time data tracking. By adopting a service-oriented approach when collecting, integrating, storing and analysing social and user data, it enabled the proposed architecture to achieve better granularity, accessibility, recommendation, and prediction of user requirements. This framework also provides service continuity capabilities that allow the system to provide value-added information that suits the user's requirements at any time and place. Having been proposed for a smart city context scenario, the framework integrates information from different sources as a necessity in real-time to make smarter decisions for individual citizens and for the city as a whole.

The study by (Satyanarayanan, 2017) gives a discussion on edge computing for situational awareness. The author defines the term *situational awareness* as an up-to-the-minute awareness that is required to move about, operate equipment, or maintain a system. This paper explores how different trends in video sensing, crowdsourcing, and edge computing can be used to create a shared real-time information system for situational awareness in a vehicular system that spans driverless and drivered vehicles. A Cloudlet in this work was used to perform video analytics of the video obtained from a vehicle, the assumption was that each Cloudlet possesses sufficient computing power and hardware acceleration in order to perform real-time video analytics on all its associated cameras. The study was conducted to help the drivers be aware of the conditions they can expect on the roads they are traveling towards. This aimed to limit road accidents. The situational awareness requires advanced analytic algorithms which will ensure sufficient speed, accuracy, and descriptive power for reporting events in real-time.

Context-awareness in this study is defined as the ability of a Cloudlet to use user browser history data to provide services of preference to the data owner. The user's browser history is referred to as small data because it defines the interests of users and can be used to provide customised online services. This caching of the historical data is to improve the performance of a Cloudlet, the QoE of Cloudlet consumers, the efficiency of Cloudlet resources, reduce Cloudlet operational cost and ensure service guarantees for consumers. The extraction of the browser history data from the mobile devices is enforced through the use of an SLA, issued in the form of a disclaimer page to ensure the benefits outlined above. One of the Cloudlet commercial products called a Xiaomi smart router has built-in abilities to capture user browsing history, using the same concept this study discusses how this function can be used to provide fairness amongst role-players and control the access and relocation of Cloudlet resources to Cloudlet consumers.

3.3 SLA-Based business models

The use of an SLA in a business model is essential in ensuring service guarantees. The aim of the SLA is to define the set of Quality of Service Guarantees, the service provider is willing to provide and at what price the consumers will be willing to pay. One of the uses of an SLA involves the management of resource allocation in high-performance Cloud auditing environments (Xiong, and Chen, 2015) where QoS metrics such as trust-worthiness, response time and availability are considered. Below are some of the SLA-based models proposed in ensuring service guarantees in a system (Casola et al., 2013) proposed an SLA-based Cloud sensing brokering platform that runs on a Cloud application consuming resources from different Cloud Providers. The platform helps form a dedicated management approach for Cloud sensors using an SLA-based approach. The SLA in the proposed work act as a guideline as to how the interactions of all role-players (customers and network providers) during the negotiation process should occur at the Cloud Brokering Platform. The network provider, also known as sensor networks, offer their resources to a Cloud application that runs independently from Cloud Providers. The sensor network offered as a service enables the users to negotiate with a provider that fulfils their desires in terms of operational and non-functional features. The issue with the proposed platform in (Casola et al., 2013) is that it lacks efficient user management and security (data privacy) strategies.

(Badidi, 2016) introduced a framework for SLA-based provisioning that relies on the broker agent to address issues such as service selection, negotiation of SLA and SLA-compliance monitoring techniques. This framework relies on Software as a Service (SaaS) to intervene between service consumers and SaaS providers, selects suitable SaaS providers that meet consumers' expectations, and negotiate the terms of an SLA on behalf of consumers. The authors also introduced a SaaS selection algorithm and decision negotiation model, which can be implemented by the SaaS broker in order to evaluate the offers of the selected service provider. Both availability and response time were used as metrics to evaluate the effectiveness in the work presented by Badidi, 2016. No prototype for the proposed work exists.

The contribution produced by the above literature assists in gaining state-of-the-art knowledge on how the SLA-based approach can be structured for a CBM taking into account the components that can be used to guide its processes and execution. Furthermore, how an SLA can be emphasised to ensure consumer service guarantees and fairness to all role-players. An important aspect of an SLA is a cost function, the following section discusses on cost functions used in a Cloudlet environment.

3.4 Business Model Feasibility

The conduction of a business model feasibility study before deployment is a crucial process required to avoid deployment failures (Maresova et al., 2017). One of the most common tools discussed in the literature that is used to prove the feasibility of a model is a Cost-Benefit Analysis (CBA). A CBA is an approach businesses use to analyse decisions, it is done through identifying the benefits a decision can bring and the costs associated with taking that decision (Maresova et al., 2017).

A CBA in this study is used to determine the feasibility of deploying a Cloudlet Business Model (CBM) by Small, and Medium Enterprises (SMEs). The functionality of a CBA is to help SMEs decision makers to take an informed decision on whether to invest in a Cloudlet technology or not (Quah, and Haldane, 2007, Maresova et al., 2017, Jenkins, 1999). The decision takes into consideration the merits of the technology such as improved QoE when accessing Cloud services through a non-monetary, customised service, low latency, high bandwidth one-hop Wi-Fi connection for customers. In general, the tool is used for the two below listed reasons (Taylor, 2015, Quah, and Haldane, 2007):

- To determine if the project is feasible, whether it is a good investment, or
- To compare one project investment with other competing projects in order to determine which is more feasible.

Below are some of the studies presented in the literature that uses the CBA to determine the feasibility of deploying a new service in a business. One study involves the usage of a CBA to help Georgetown hospital decision makers to decide whether or not to retrofit the hospital. The decision to retrofit can help turn the hospital into a smart one through the addition new technologies in the hospital to improve health care service delivery such as ventilation, security, hygiene, accessibility, conservation, lighting, sanitation, and aesthetics. Two options were considered for conducting the analysis either *do nothing* or *retrofitting for smarting the hospital*. The results showed that the *do nothing* option will not provide any benefits to the medical facility (Quah, and Haldane, 2007).

A model to assess cloud computing with respect to its utilisation in business practice for assessing the efficiency of investment was developed. This model uses a multiple-criteria analysis of variants such as systematic literature review, analysis of real cloud computing services, expert interview, and a case study was applied in order to develop and evaluate the model. The CBA was then used to evaluate and assess the efficiency of cloud computing adoption in an organisation. The results of a CBA indicated that the model suits the practical and decision-making requirements in cloud computing (Maresova et al., 2017).

In some cases, businesses deploy the proposed services even though the Net Profit Value (NPV) and the Return of Investment indicators do not show favourable results. An example includes the deployment of Cloud Computing in the Czech Republic, where it was deployed although results showed that they can save up to 33% of operating costs per year with the return on investment over three years (Maresova, and Puzova, 2014). However, the authors elaborated that the unfavourable results were caused by the lack of including qualitative benefits due to the unwillingness of the organisation to share data. In summary to conduct an effective CBM that can enable informative decision making requires the consideration of both qualitative and quantitative costs.

This study uses the first reason mentioned above for using a CBA to determine whether deploying a Cloudlet is a good investment or not for SMEs. To determine the feasibility of deploying a Cloudlet three steps are followed: identification of costs, benefits, and lastly, comparing costs and benefits.

3.5 Cost Functions used in a Cloudlet environment

A cost function known as a price model is one of the metrics that must be included in SLA negotiations and is defined by the service provider. It defines how actions of the role-players will be captured in the agreement. For example for consumers to gain access to a service, they must be willing to pay a certain price (Xiong, and Chen, 2015, Patel et al., 2009). Similarly, for service providers to gain more consumers they must be willing to spend more on improving their service delivery and quality. Furthermore, the cost function also considers the compensation of SLA violations.

One of the main differences between the Cloud and the Cloudlet is the availability of resources, in the Cloudlet resources are limited but in the Cloud they are unlimited. Therefore, in the case of Cloudlet, the competition-based price functions are effectively used for resource allocation to mobile devices in the Cloudlet environment. Three cases of competition based pricing models are discussed below namely, the auction, non-cooperative game, and supply-demand pricing models.

3.5.1 Auction-based pricing

A real-time group-buying auction is used in a Cloudlet that offers its services (mobile videos) to mobile devices in its proximity with lower prices while maximising the profit of the group (Zhang et al., 2014). In this type of auction, the seller creates a price curve of auction prices, and the buyers with a price higher than the auction price are the winners. The buyers get more discount if more buyers participate in the auction. The merits of the model are that it supports pricing minimisation and expected profit maximisation. The drawback is that it does not guarantee strong truthfulness, which can cause conflicts amongst the parties involved in an auction. This drawback can be caused by buyers placing multiple bids using false names in order to receive a higher discount.

The lack of guaranteed truthfulness was solved in (Jin et al., 2016). The authors addressed this problem by charging the mobile users according to the payment policy of the Vickrey auction. The model consists of three stakeholders, namely the mobile users, Cloudlets and auctioneer. The Cloudlet serves (processing, storage, and networking) only mobile devices in its proximity and the auctioneer sorts the mobile users (buyers) in ascending order of bids and descending order of tasks, in order to select a winning buyer and seller. In cases, where the buyer wins more than one seller, the auctioneer selects only one seller where the buyer's utility is the highest. Although, false name bidding can be experienced with this model it also provides individual rationality, budget balance, and truthfulness which is beneficial for mobile users.

3.5.2 **Non-cooperative game pricing**

The non-cooperative game (Jin et al., 2016) has been widely used for resource management in Cloud networking, for example amongst brokers in a Cloudlet environment in order to achieve high profit for Cloud Providers. This model is used in cases where multiple brokers exist for providing Cloud resources and the brokers are said to be selfish. Each broker in this model cares only about maximising their profit payoff, without being concerned about other brokers nor the welfare of the network. The Jacobi best-response algorithm is used in this model to achieve an approximation of the Nash equilibrium at which the bid prices of all brokers are optimal, this helps minimise costs for the mobile users.

3.5.3 Demand-Supply pricing

The work by (Wu, and Ying, 2015) introduced a Cloudlet-based multi-lateral resource exchange framework for mobile users. In this framework, a demand-supply pricing model is used, where each mobile user owns a Cloudlet and acts as a buyer and a seller. The aim of each individual mobile user is to increase their payoff which is the difference between the utility of buying resources and the cost of selling those resources. The mobile users use an aggregate excess demand function to define the total demand and supply. The merits of this pricing model are that it maximises payoff and improves resource efficiency. The author (Tiwary et al., 2017), introduces a Cyber foraging pricing scheme (CPS) that declares real-time prices for mobile users based on consumption of resources during the peak-hours which is dependent on the demand-supply curve. A CPS is a dynamic, distributed real-time pricing scheme for cyber foraging systems, which behaves fairly with home and foreign users. In this work, the home users are classified as users in the specific location where the cloudlet is deployed and foreign users are the users around the Cloudlet coverage area. The scheme offers optimal prices during the peak hours and broadcasts the prices to the mobile users, enabling them to select an appropriate Cloudlet during peak hours which will minimise their costs with reasonable SLA compliance. The drawback about this work is that the pricing scheme is based on only the proximity of the Cloudlet, not its mobility. The consideration of mobility of a Cloudlet can help to increase the QoE for Cloudlet users.

A third party is introduced in (Panigrahi et al., 2015) to sell Cloud services on behalf of the Cloud Provider. In the Cloudlet as Agents of Cloud Brokers (CAB) architecture, the Cloudlet is placed nearer the users in order to help the brokers to solve computational intensive services with strict SLAs for the user. In the pricing scheme introduced in the CAB architecture, the broker charges for the brokerage cost as well as priority return of results by the Cloudlet; the charge is termed a CLT charge (Panigrahi et al., 2015). In cases, where the Cloudlet completes a request without any interruptions or a service whose deadline is met as written in the SLA, the broker charges the CLT charge in-full. In other instances where a service is switched due to profit optimisation (the SLA is not met), the broker charges less than full CLT by selecting a value on a descending scale containing pro-rated CLT charges. Lastly, if the response time is greater than or equal to that of a remote Cloud, no charges are filled by the broker. The results obtained from the simulation of the architecture and its pricing model indicates that to optimise the profit, the Cloud broker should deploy the Cloudlet near to its users. The authors focus on how the brokering agent in a CAB architecture can optimise its profit not how the Cloudlet users can benefit from its existence.

This work is focused on developing a Cloudlet Business model for SMEs such as coffee shops and community shopping centres. This model uses an appropriate pricing model which adheres strictly to the agreed-upon SLA, promote fairness, and maximizes profit for service providers. The pricing models considered in this study are the demand-supply and fixed pricing models. To gain access to consume Cloudlet resources, the Cloudlet consumers are expected to share their browser history data. The browser history data is considered as a fixed non-monetary value a Cloudlet consumer pays for a session with the Cloudlet. Due to the nature of compensating the provision of Cloudlet resources, the Cloudlet owner is responsible for the Cloudlet operational costs. For the resources used by Cloudlet consumers, the Cloudlet owners use a monetary value (Demand-supply) to pay Cloud or Internet Service Providers, depending on the demanded and distributed resources on the network. Fairness is achieved in the sense that as Cloudlet consumers share their context, the caching of this data minimises power consumption costs incurred by the Cloudlet owner to keep a Cloudlet in operation. Service level agreements are mechanisms for capturing pricing models in this work, hence the effect of SLAs on Business Models is discussed next.

3.6 Discovering the effect of an SLA on a business model through a simulation method

Several works have been published pertaining to the simulation of a Cloud environment to validate the SLA. The study conducted by (Khattak, and Jalal, 2015), proposed an SLA performance-based framework which extends the WSLA. The framework converts low-level metrics to high-level, in order to measure them and analyse the performance of the Cloud system against the SLA agreed upon by the parties. The validation of this framework was carried out on CloudSim tool, and part of the simulation setup considered different VM policies such as Maximize Throughput Provision Policy (MTPP), Minimize Response time Provision Policy (MRPP) and Maximize Utilisation Policy (MUPP). Based on the results the MUPP performed better in terms of deadline satisfaction as the tasks increased in a network. The simulation was carried out to show the functionality of SLAs across VM policies, price mechanisms, and resource allocation. To improve the resource provisioning, an SLA violation detection mechanism was proposed to help in allocating efficient processors for task processing (Musa et al., 2016). The assumption carried out by the mechanism is that the VM allocated by the broker agent should be based on the tasks initiated (either greater than or equal to the number of tasks to be processed). This was carried out on Cloudsim tool where the authors considered two simulation setup scenarios: In the first scenario, the number of resources (virtual machines) is greater than the number of resources requested by the submitted jobs and in the second scenario, the number of provisioned resources is less than the number of resources requested.

The first scenario shows the success of the SLA as the resources allocated are sufficient to render the services requested for processing. Whereas, scenario two shows violations of the SLA as the resources needed to complete the execution of requests are not met by the service provider. However, the work lacks detailed information about what happens when the SLA is violated. Another work similar to the one done by (Musa et al., 2016) proposed an approach to monitor SLA compliance implemented at the clients' end (Suneel, and Guruprasad, 2015). The monitoring of SLA compliance as clients initiate requests for processing on the Cloud, included first step where a client generates an information fetch task using a Gen (SLA) function, where the SLA is considered as an input to the function. Secondly, the Gen (SLA) function generates a task that consists of instructions to collect relevant information regarding SLA compliance status in the Cloud. Thirdly, the information fetch task is forwarded along with a set of tasks for processing to the Cloud instance. Lastly, when the task arrives in the Cloud, a log is created. The log documents the arrival time and hash of the tasks. The Cloud then executes the tasks and relays the SLA status to the provider. Although the authors mentioned that they implemented their proposed algorithm on the CloudSim tool, the research work lacks results to show proof of concept of their work. Also, this work sends feedback of the SLA status to the provider but lacks actionable policies in the case of an SLA violation.

The simulation carried out in (Suneel, and Guruprasad, 2015) is similar to this study. Although, both studies involve notifying the service provider regarding SLA violation and possible issue/s that might have caused it. This study further provides a penalty that a service providers can provision when SLA is breached depending on the task expected completion time, the user session will be extended.

3.7 Chapter Summary

This chapter speaks to the first two objectives of the study that brings about the outcome of the third objective. Objective One: To investigate existing SLA-based Cloud Business Models or similar models that can be aligned to Cloudlet environments and a price function that fits such environments. Objective Two: To explore how the use of small data sensors can be beneficial for an SLA-based Cloudlet business model. The chapter addressed the above-mentioned objectives by discussing the literature on adoption of an SLA and its functionality on both Grid and Cloud computing. Lastly, the mechanisms used to enable context-awareness on the Cloudlet and how it brings about value were covered. The knowledge gained in this chapter through the literature provides a structure that can be leveraged for the development of an SLA-based Cloudlet business model discussed in the following chapter.

Chapter 4

Service Level Agreement (SLA) Based Cloudlet Business Model Development

4.1 Introduction

A business model can be defined as the value logic of a business in terms of how it creates and captures the value offered to customers and is represented by an interrelated set of elements that address the customer value proposition, business architecture (Key Resources and Processes) and the economic dimensions (Christensen , and Johnson, 2009). A business model can be either beneficial to the service providers (top-down) or service consumers (bottom-up). As indicated in chapter one, the goal of this research work is to develop an SLA-based Cloudlet Business Model (CBM) that would be beneficial to all parties involved. Therefore, the CBM uses a bottom-up business model approach in order to help facilitate the deployment of a Cloudet by SMEs. It outlines the impact and interactions of components to bring value and benefits related to the model role-players. This chapter centres on the important aspects that build up the CBM such as the processes and role-players. Included in the discussion is the importance of small data and SLAs in the CBM including, how the SLA is enforced in the network. Lastly, it uses the Four Box Model to describe the value logic of the CBM.

4.2 The SLA-Based Cloudlet Business Model Processes

This section discusses the processes of the Cloudlet Business Model (CBM). The model processes such as job initiation, data extraction, data analysis, and job delivery can help ensure service guarantees on the network. The events in the processes and the

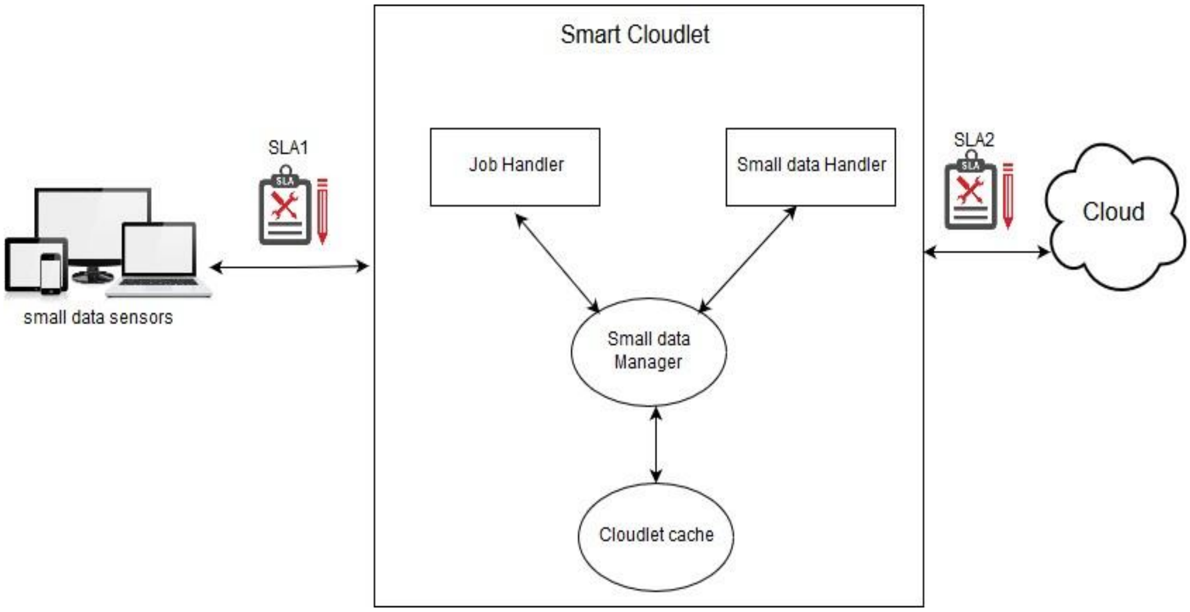

Figure 4.1: A Cloudlet Business Model

communication between the components of the CBM are carried out in the discussion below with reference to Fig. 4.1. The diagram in Fig. 4.2 illustrates a simplified flow of events as Cloudlet consumers initiate tasks for processing to the network. It shows the occurrence of events that happen on the network triggered by certain conditions that must be met for Cloudlet consumers to achieve a better QoE. The events in Fig. 4.2 are further explained in detail the below under listed CBM processes except for the events of the Cloud (where a Cloudlet consumer did not share data).

4.2.1 Job Initiation

A mobile user (SME customer) initiates a task to a Cloudlet for processing. To grant a task initiator access to a Cloudlet, the job handler (shown in Fig. 4.1) calls the SLA executor component to take care of the logistics. The SLA executor issues a captive portal that requires the initiator to agree to the conditions of service to allow the use of Cloudlet resources. The type of captive portal used is a disclaimer page that details the legal responsibilities between the SME customers and the Cloudlet owner, which an SME customer must accept if they require access to a Cloudlet. The captive portal is configured on the network interface of the deployed Cloudlet access point/ router by the Cloudlet owner.

Among the conditions that must be stated on the captive portal, is enabling the extraction of browser history data from the mobile device to the Cloudlet for cache purposes. The sharing of historical browser data means that the consumer agrees to the SLA (SLA1) terms given by the Cloudlet owner. SLA1 is a label given to the SLA between Cloudlet consumer and owner. After accepting the conditions, the SLA executor notifies the SLA monitor about the request status and the task is further processed to the data extraction stage. In the case where data extraction did not occur, the tasks are transferred for processing to the Cloud as shown in the flow chart. The SLA monitor is responsible for verifying the SLA agreement and overseeing that the stated Service Level Objectives (SLOs) are met.

4.2.2 Data Extraction

After the Cloudlet consumer has agreed to share the browser history data, the Cloudlet extracts and cache the browser history data to its cache. The Cloudlet pushes the data to the Cloud for storage purpose and pre-fetches it in a case when the data is requested by the edge. Before the data is cached, the task requested or processing is matched against the content of the Cloudlet cache for existence.

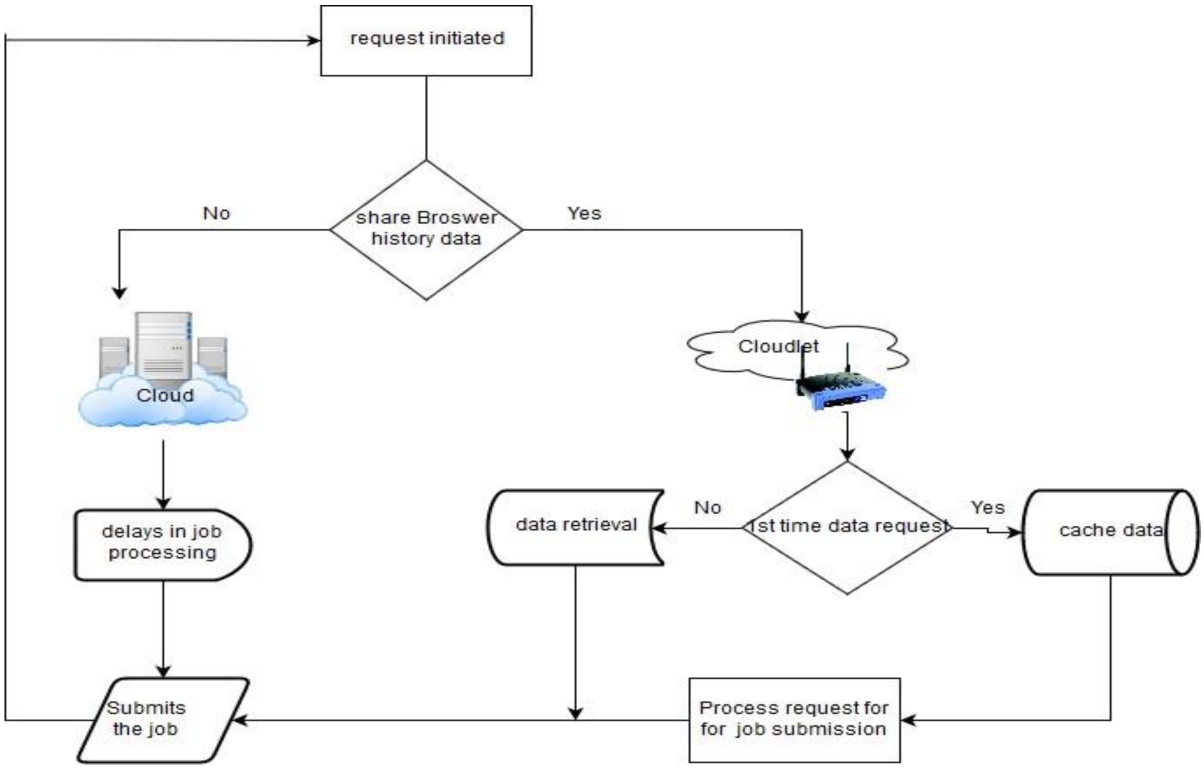

Figure 4.2: A flow chart of the CBM processes

This is performed to avoid data redundancy. If the job of the requested task exists in the cache, it is pre-fetched from the Cloud or Cloudlet to the Cloudlet consumer, meaning the data analysis process or resource allocation is skipped to the job delivery process. Otherwise, the task is forwarded to the next stage for analysis. Within the job handler, the small data handler is responsible for maintaining the connection between the Cloudlet and the small data sensor for the successful extraction of mobile users' browsing history. While the job handler is responsible for taking care of the task requested for processing by allocating the appropriate resources, scheduling, and access policies.

4.2.3 Data Analysis

In cases where the task requested for processing is not matched against any content in the Cloudlet cache, the task is forwarded to this component for processing to find an appropriate job. The job handler component allocates resources for processing the task. The task can be sent to the Cloud for processing or processed in the Cloudlet depending on its size. The most commonly used technique in Cloud computing to analyse data is called Hadoop (M., and R., 2014).

Although Hadoop is commonly known to analyse big data in the Cloud, it can also be used for analysing small data (Cloudera, 2014, Mor, 2014). Using a one-hop connection, the Cloudlet forwards the task for processing to the Cloud and the job is forwarded back to the Cloudlet by the Cloud. The job query processed in the Cloud is concatenated with previously stored job requests in the Cloudlet cache, this is handled by the small data manager. The use of Hadoop will speed up the data analysis, processing and decrease the job duration. A small data manager is responsible for keeping track of a Cloudlet's consumer history of activities and stores them in a Cloudlet cache. This component uses a data replacement algorithm to ensure that the data stored in the Cloudlet cache is useful for consumers. It achieves the above by calculating the ranking units of same data requested in the Cloudlet for processing by other edge users (Zhijun et al., 2017). This implies that even first-time Cloudlet consumers can enjoy preferential services.

4.2.4 Job Delivery

After the job scheduler has allocated appropriate resources to be used for the request, the job handler sends the output of the request back to the initiator (mobile device).

The SLA monitor in this instance comes into play to check if the processing and the allocation of the resources match the objectives stated in the SLA. The SLA in this context is considered as SLA1 depicted in Fig. 4.1 which is an agreement between Cloudlet users and the Cloudlet owner as agreed to by the role-players; if not what are the penalties involved to ensure fairness? This rule is based on the observations captured from the use of algorithms or third-party monitoring systems such as Cloudlet Control and Monitoring System (C2MS) (McGilvary et al., 2013).

4.3 The role-players of the Cloudlet Business Model (CBM)

The CBM is composed of three role-players namely, the Cloudlet consumer, Cloudlet owner and the Cloud Providers or Internet Service Providers.

4.3.1 The Cloudlet consumer

The term Cloudlet consumer, mobile user, SME customer or small data sensor owner all refer to an SME customer that is willing to share their browsing history in order to gain access to using Cloudlet resources for Internet connectivity. The process of sharing is guided by the SLA (SLA1) that is issued in the form of a captive portal when a request for processing is initiated. The use of an SLA is to enforce Cloudlet access policies and resource allocation for ensuring customer QoE by Cloudlet owners. The ensuring of QoE by the Cloudlet owner enables this role player to gain interest in the business, and share information with other potential consumers which will later improve the SMEs Quality of Business (QoB).

4.3.2 Cloudlet owner or SME policymaker

A decision maker or influencer in an SME such as a coffee shop, doctor's office, Internet cafe, library or any other SME can be a Cloudlet owner. The Cloudlet owner takes an informed decision to deploy Cloudlet technology as a means of attracting more customers in order to improve their marketing strategies, QoB and revenue generation. In ensuring service guarantees and QoE the Cloudlet owner leases resources from an ISP/ Cloud Service Provider. The Cloudlet owner might lease resources and equipment from both the Cloud Provider and ISP or might lease only from the ISP. The leased resources determine the QoB a Cloudlet deployment will add to an SME and is measured by the customers' QoE. The Cloudlet owner is also responsible for enforcing Cloudlet usage policies. The policies are enforced by configuring a Cloudlet (example: Xiaomi smart router) in a network layer to issue a disclaimer page when a request is initiated for processing.

4.3.3 Cloud Provider or Internet Service Provider (ISP)

The Cloudlet owners may use both Internet Service Provider (ISP) and Cloud Providers for leasing resources to ensure an effective operation from a deployed Cloudlet. The ISP enables Internet connectivity by providing Internet bandwidth packages, access points or routers, and equipment configurations. The Cloud Providers namely, Google, Microsoft, and IBM enable Cloudlet owners to lease resources such as processing power and storage capacity using a pay-as-you-use price function. The relationship between the Cloudlet owner and Cloudlet resource providers is controlled by an SLA (SLA2) that defines the conditions of their partnership in terms of the service offered as measured using either a dynamic or fixed price function.

Fig. 4.3 shows the interaction between the role-players in a business model. The three most important entities in strengthening the relationship between these role-players as shown in Fig. 4.3 are Quality of Experience (QoE), Quality of Service (QoS), and Quality of Business (QoB). The QoE is measured by the connection delays, throughput and service fees. The QoB is measured by the ability of Cloudlet owners to provide QoE by ensuring availability of a Cloudlet and its resources. Lastly, the QoS is measured by the fulfilment of SLA stated metrics such as task completion time and waiting time by the Cloudlet owner.

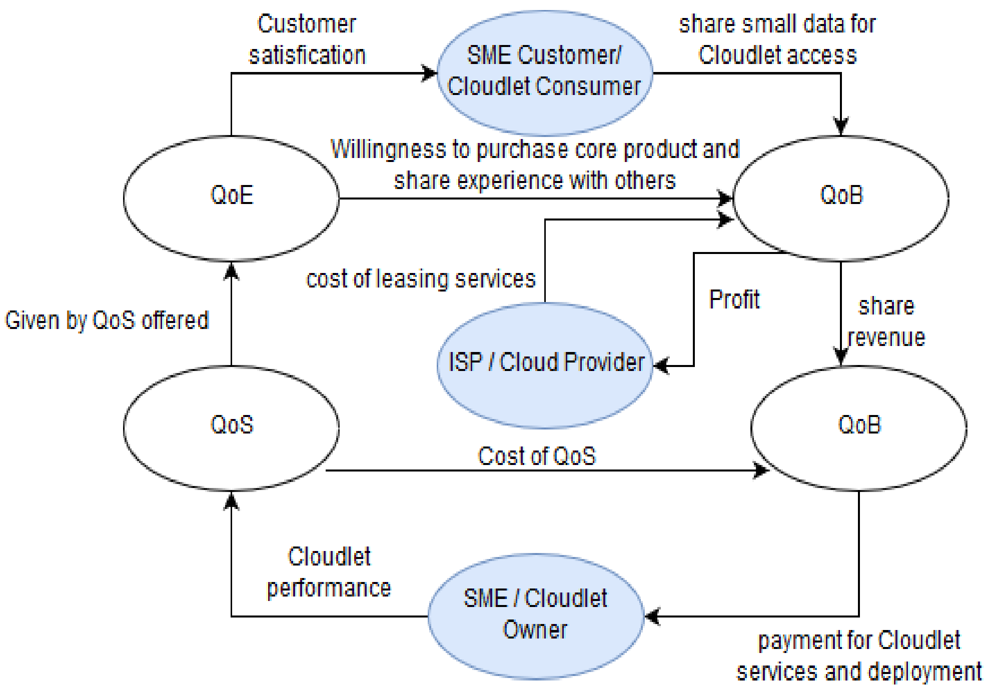

Figure 4.3: Business ecosystem for role-players in the CBM value chain

4.4 CBM use-case

A CBM can be applied in businesses that are small in size situated in busy areas such as in shopping centers, malls and others, willing to explore ways in which they can improve customer satisfaction and attraction. This work pays close attention to the Coffee shop business. Coffee shop businesses are usually small in size and located around busy business areas in big towns or cities. These intend to target employees working in the surrounding environment and other customers. The peak hours for coffee consumption are usually during the early mornings when employees are traveling to work, students going to campuses and during their lunch times. An assumption is that a coffee shop sells an average of 250 cups of espresso and coffee a day to its consumers (E-Imports, 2017). Deploying a Cloudlet in such an environment may indirectly attract the SME customers whose clients spend more hours on the premises due to better performance of a one-hop, low latency, non-monetary Wi-Fi service. It can also, extend the peak hours or times when visitors are coming in to consume value-added services (beyond espresso and coffee). Among the customers will be start-up SMEs who would want to utilise the resources/services offered via a deployed Cloudlet to run their businesses. In this setup, the coffee shop provides for an SME renting purpose-made space in its premises. This will provide flexibility plus exposure for the coffee shop and growth for start-ups. In this case study, the Cloudlet consumers are coffee shop visitors both regular coffee consumers or short-term visitors, and the SME owners and their clients or longer-term visitors. The Cloudlet owner and the owner of the coffee shop are assumed to be the same person or entity.

In the event that a Cloudlet instance is already in operation and the SMEs on the premises are offered service packages as incentives to attract Clients to the coffee shop, the success of the deployment relies on the leased resources such as Internet service packages from the ISP and the existence of efficient resources such as Virtual Machines (VMs) to cater to the needs of visiting clients. Furthermore, the deployment success parameters should take into consideration the peak hours, when a number of requests may be initiated to the Cloudlet. Fig. 4.3 shows the factors that influence the success of deploying a Cloudlet and the benefits that come with them. As shown in Fig. 4.3 the QoS is maintained by the coffee shop owner by providing efficient yet effective resources to improve Cloudlet performance. This is achieved by the lease of efficient resources from service providers. The higher the QoS in the network, the more consumers are exposed to QoE which influences them to choose the coffee shop for their future visits. This improves the coffee shops' QoB, which can further improve both profit and business marketing.

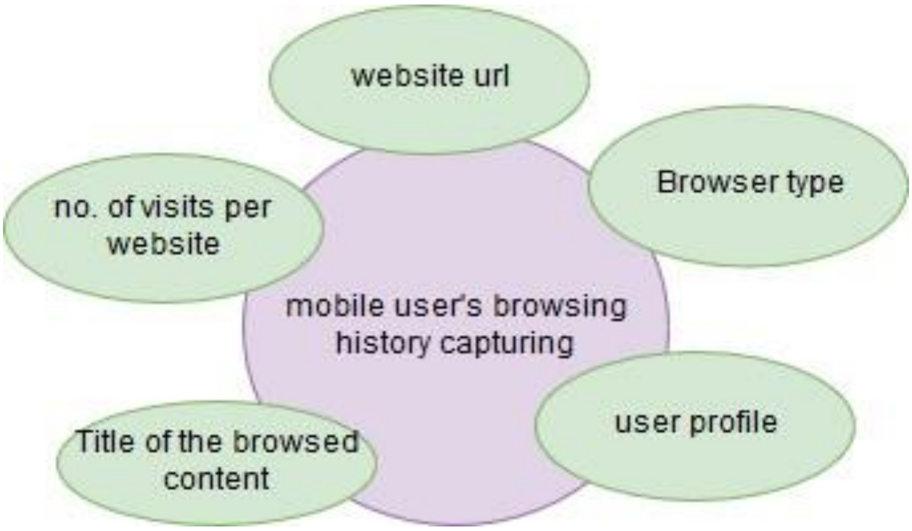

Figure 4.4: Parameters of a browser history log file

4.5 Data cache for compensating the Cloudlet resource consumption

In ensuring the service guarantees and improvement of Cloudlet consumers' QoE, small data is used to leverage the capabilities of a Cloudlet by enabling a Cloudlet to provide preferred services on the edge. Small data refer to the browser history data that defines the behaviour and browsing patterns of individuals. Capturing browser history data is already achieved by one of the Cloudlet commercial products called the Xioami Smart Router (Pang et al., 2015). The browser history data is obtained from the mobile browser historical log file. The log file is generated and updated as the user continuously browsing the internet. It consists of information about the visited web pages (URL, keywords used for searching, timestamp, user profile, and others) as shown in Fig. 4.4. A Cloudlet being aware of the edge user's preferences enables it to predict future edge user access. The prediction of future preferences uses a collaborative filtering algorithm which connects the interests of different individuals to build an effective prediction model to predict future behaviour of data owners (Herlocker et al., 2000). This means that even for first time Cloudlet consumers the Cloudlet will recommend services that might be of interest to them and this is adjusted as more data is shared by Cloudlet consumers.

The use of small data in the CBM brings about advantages such as customised service provisioning, minimised task completion time, low Cloudlet power usage because of the reduced processing cycles (efficient resource management) and minimum cost for leasing resources (reduced Cloudlet operational cost)

It is evident that customers do not pay a direct usage fee for Wi-Fi services. However, using the shared browser historical data to provide customised online services can help minimize the cost to the paying role-players, reduces the Cloudlet resource usage of a Cloudlet when processing initiated requests (reducing power costs and the cost of additional resources), and improve consumer QoE. The sharing of the users' browser history data follows a fixed non-monetary price function which implies that Cloudlet consumers pay less for bulk usage of Cloudlet resources. The fixed non-monetary price function is determined by the Cloudlet owner as stated on the SLA (SLA1) to attract the Cloudlet consumer to choose their premises as a hub for Internet usage. Furthermore, sharing of the browsing history is carried out through accepting the terms given on a captive portal. Through sharing their browser history an SME customer can enjoy a one-hop, resource-rich, low latency, non-monetary Wi-Fi connection.

4.6 Service Level Agreement for ensuring service guarantees and customer Quality of Experience

An SLA describes the offered service, guarantees and penalties involved between the agreement of a resource provider and consumer (Khattak, and Jalal, 2015), Serrano et al. (2016), Patel et al., 2009). In Fig. 4.1 (Cloudlet Business Model) there are two instances of an SLA presented. SLA1 is issued between the Cloudlet consumers and the cloudlet providers. On the other hand, SLA2 is issued between the Cloudlet owner and Cloud Provider or ISP, the Cloudlet owner uses this SLA to ensure QoS (Middleton et al., 2009). SLA2 has a direct impact on the outcome of SLA1, for instance in a case where an ISP encounters a systematic problem it can affect the distribution of their service to Cloudlet owners, resulting in SLA1 violations due to the network outage. The above is summarised by Fig. 4.3. Therefore, the use of bipartite SLAs (SLA1 and SLA2) in this work is to achieve QoE (that encapsulates delays and service fees) and service guarantees for the Cloudlet consumers' requested tasks for processing. The two SLAs are mapped using different price models: SLA1 is mapped using a fixed non-monetary value which is a small data and SLA2 uses a dynamic monetary value agreed by the involved parties.

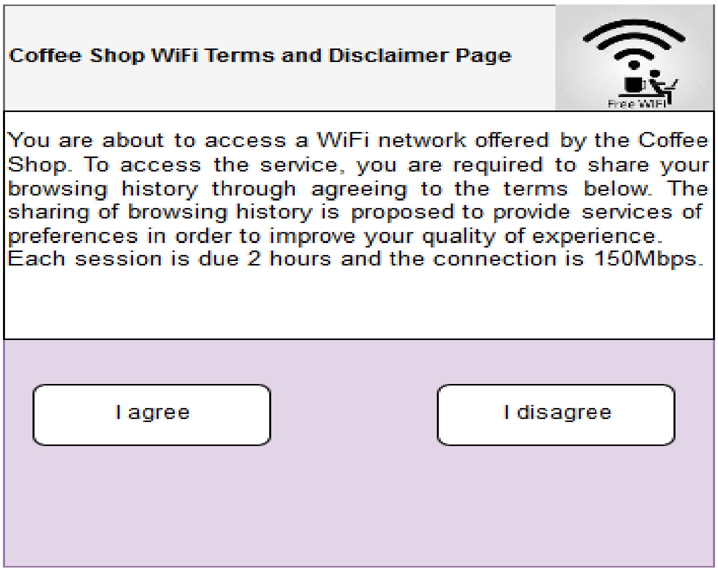

Figure 4.5: An Example of a Disclaimer page an SME can use to enforce Cloudlet access policies

4.7 Enforcement of an SLA on Cloudlet Consumers'

The enforcement of a Cloudlet access policy is carried out through the usage of a captive portal called a disclaimer page as shown in Fig. 4.5. The captive portal is configured on the network interface of the installed router/ access point by the Wi-Fi Cloudlet owner (Belton et al., 2015). A disclaimer page serves as an SLA1 and comes in the form of a pop-up page on a Cloudlet consumers' device when they initiate a task for processing. As shown in Fig. 4.5, it consists of a statement that outlines the legal responsibilities, session duration, and bandwidth allocation to which an SME customer must agree before proceeding (Purdy, 2011). Also, captured on the disclaimer page is the consent to allow the capturing of an SME customers' browsing history for Cloudlet use to improve its QoE. In a case where an SME customer does not agree to the terms stated on the disclaimer page the user is denied access to the Cloudlet resources.

4.8 The value proposition of a Cloudlet Business Model

A successful business model is a model that is able to address the value logic by indicating what the value is, and how it will be captured and compensated to ensure fairness to all role-players involved in the business model (Wollenick et al., 2012, Padilla et al., 2015). In the CBM the above is achieved by using a Four Box Model (shown in Fig. 4.6). A Four Box model is used to summarise the CBM by simply highlighting the interaction of the components on the CBM in bringing out the business value logic, and benefits involved (Fielt, 2014). The model consists of four elements namely customer value proposition, profit formula, key resources and key processes that are explained as follows (Fielt, 2014).

4.8.1 Customer Value Proposition (CVP)

The CVP is defined as a customer-oriented problem that a Cloudlet deployment in an SME environment can help solve. The problem that the deployment of a Cloudlet instance will help optimise is the poor QoE presently encountered by mobile users when using Cloud-Wi-Fi to access distant Cloud services. This problem is solved at the expense of mobile users sharing their browsing history. By deploying a one-hop, low-latency, resource-rich, non-monetary Wi-Fi connection in the scenario given above, a business can attract more customers, improve their marketing and purchasing patterns of services offered, as shown in Fig. 4.6.

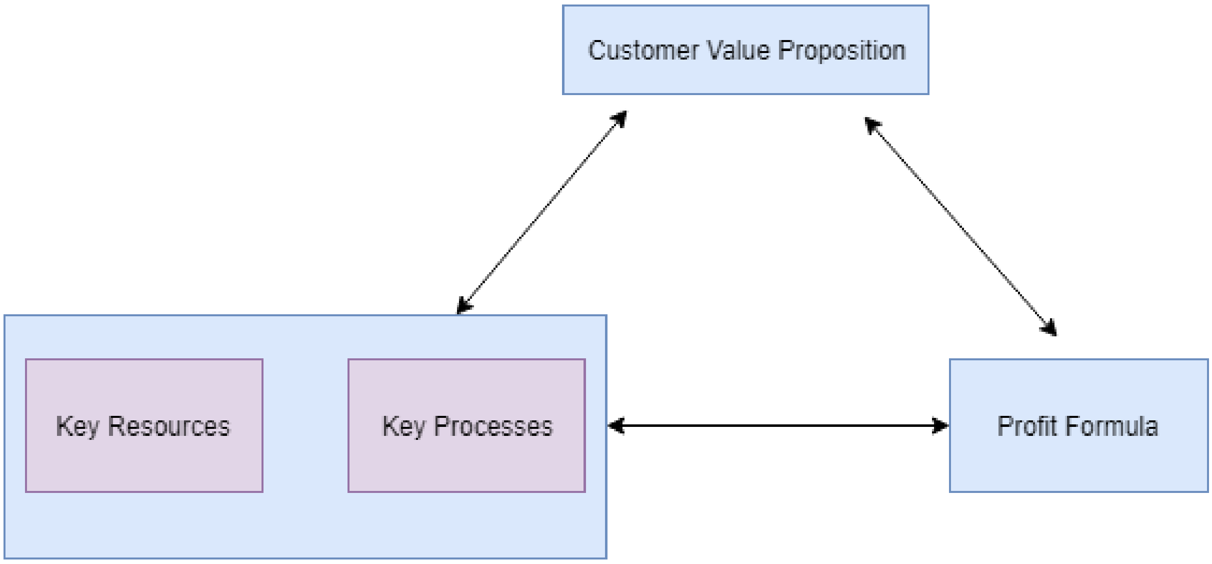

Figure 4.6: A Four Box Value Proposition Model

4.8.2 Profit Function

Defines how the access to a Cloudlet will be compensated, and who will pay the cost of resources required for the Cloudlet to be operational. Due to the nature of the service offered the customers' do not pay a monetary value for the usage of Cloudlet resources. However, in ensuring fairness to the role-players responsible for the cost of resources, a mechanism (small data sharing) to reduce the high consumption of resources is used. This mechanism is enforced to efficiently save Cloudlet resources by caching mobile users browsing history (fixed cost) to offer services of preference, decreasing the cost of power and additional resources that might be needed. The Cloudlet owners are responsible for leasing resources or equipment that will ensure service guarantees and QoE from ISPs or Cloud Providers using a dynamic price function (demand-supply price function), depending on the agreement. The price function is captured on the SLA between the parties in agreement.

4.8.3 Key Resources

This element outlines the entities that are needed to deploy a Cloudlet instance in an SME environment. These entities are not limited to Cloud Providers, equipment, ISPs or technology companies that will provide Internet service packages, installation, and configuring of devices such as a router. Through this SMEs such as a coffee shop can obtain a QoB which is measured by the quality of resources leased, service offered and the QoE encountered by end users. However, to avoid paying more for less a dynamic price function is more suitable for the leasing of resources and software by the Cloudlet owner from an ISP or Cloud Providers.

4.8.4 Key Processes

This model uses an SLA as a policy to facilitate and ensure a successful, yet fair deployment of a Cloudlet business model by SMEs. The SLA is used as a mechanism to improve the service offered to end users by ensuring that resources in place are efficient for all requests initiated. In cases where an SLA is violated, in this study, the extension of Cloudlet consumer session duration is considered as a way to improve QoE when violations occur.

The explicit description of the four parameters of the Four Box Model plays a significant part in bringing out the value of deploying a CBM in a coffee shop environment and the benefits involved. This is further discussed in details in Chapter 4 using a Cost-Benefit

Analysis to prove the feasibility of the model.

4.9 Chapter Summary

This Chapter provided the description of the processes involved in developing a CBM. Outlined in this chapter include the model's important factors such as role-players, processes, use-case scenario, value logic and the function of the SLA in the model. This chapter addresses objective three of the study that reads as follows: To design a proposed Cloudlet business model that integrates the SLAs of role-players with an appropriate price function. The following chapter proves the feasibility of the developed Cloudlet Business Model using a coffee shop scenario.

Chapter 5

Feasibility of the Cloudlet Business Model

5.1 Introduction

The processes involved in the development of a Cloudlet Business Model was discussed in the previous chapter, this included the discussion of an SLA as a means of minimizing the Cloudlet operational costs as well as processes involved in ensuring the efficient deployment of a CBM by the SMEs. This chapter focuses on demonstrating and discussing the benefits SMEs can leverage from the deployment of a CBM, taking into account the deployment cost and the outcomes that the deployment can yield for SMEs. Through using the Cost-Benefit Analysis (CBA) method, the aim of this chapter is to prove the feasibility of a CBM using a coffee shop scenario as a use-case. The chapter is organised as follows: identification of costs and benefits based on role-players and the comparison of costs and benefits.

5.2 Identification of Cost and Benefits

This section determines the costs involved in deploying a Cloudlet with the benefits it can yield. The determined cost and benefits are based on CBM role-players. This is done to clearly point out the roles played by individual role-players in the CBM. Below listed are individualized costs and benefits for each CBM role-player. The CBM role-players as discussed in chapter four are Cloudlet Owner, Cloudlet Consumer and

Cloud Provider/ Internet Service Provider.

5.2.1 Costs

The costs are represented as quantifiable and non-quantifiable for developing a Cloudlet Business model, and are considered negative factors. Negative factor in this context is an act of using monetary value for an exchange of a service.

5.2.1.1 Quantifiable costs per role-player

The Cloudlet Consumer only pays monetary value for the cost of the core service or product offered by an SME.

The Cloudlet owner is responsible for the cost of leasing an Internet Service Provider (ISP) or Cloud Provider's resources, Cloudlet technology and IT infrastructure cost which is needed to ensure the deployment and operation of a Cloudlet on an SME environment. Furthermore, the Cloud Provider or Internet Service Provider (ISP) do not have quantifiable cost because they would not be paying but benefiting from the leasing of their services to Cloudlet owners.

5.2.1.2 Non-quantifiable costs per role-players

Cloudlet consumer is responsible for sharing browser history data as means of paying for an attempt to consume Cloudlet resources. The sharing of browser history is referred as a non-monetary value.

The Cloudlet owner has to take into account the risk of deploying a Cloudlet as it can yield or loose profit. If the Cloudlet owner, does not yield any benefits, the Cloud Provider or Internet Service Provider (ISP) will have to face the risk of leasing resources to Cloudlet owners.

5.2.2 Benefits

The benefits are represented as quantifiable and non-quantifiable for developing a Cloudlet Business model, and are considered positive factors. A positive factor in this context is an act of gaining monetary value for an exchange of a service.

5.2.2.2 Quantifiable Benefits per role-players

In the event of a Cloudlet deployed on a SMEs environment, the Cloudlet consumer gets to use the Cloudlet over an exchange their browsing history (non-monetary value exchange), experience and enjoy a service of interest with a fast execution rate (improved QoE), awareness of core products discounts, and higher service availability.

Improved sales and marketing of the core product and business, low IT infrastructure cost are benefits leveraged by the Cloudlet owner. With that said, the Cloud Provider or Internet Service Provider (ISP) benefits with the increased number of businesses leasing cloud services.

5.2.2.3 Non-quantifiable Benefits per role-players

The use of a browser history data as an exchange non-monetary value for consuming Cloudlet resources will benefit the Cloudlet consumer by reducing the amount of power (mobile device energy consumption) used to check new searched contents from the internet.

Benefits such as simplified service management, increased number of new customers, high service availability, enhanced business reputation, and help keep-up with competitors are among endless benefits a Cloudlet owner can receive through deploying a Cloudlet Business Model on a SME environment.

5.3 Cost Benefit Analysis (CBA)

The CBA is used in businesses as a way of summing up *benefits* or action and subtracting the *costs* associated with taking that action. The CBA is therefore used for decision making involving decision makers and it the context of this research study, the Cloudlet owner. This is because the Cloudlet owner is responsible for deciding whether or not to deploy a Cloudlet technology on a SME environment (Maresova et al., 2017). The decision takes into consideration the amount and quantity of equipment and Internet service rental from an ISP. The equipment required includes but is not limited to a Xiaomi Smart Router, Ethernet cables, modem (usually comes with a package offered by the ISP) and amplifier to extend the strength and coverage of Wi-Fi signal.

The below equations articulates the cost associated with deploying a Cloudlet on a SME environment such as deployment cost, and cost due to leasing resources from ISPs. The Table 5.1 describes the abbreviations used on the deployment cost equation.

Table 5.1: Abbreviations used in the CBM analysis

Abbreviation	Description
DC	Deployment Cost
Pr	ISP or Cloud Provider
K	Number of VMs
C	Cloudlet router fixed cost
E	Equipment such as Internet cables, modem, and others.
Fr	dynamic Pr resource cost for leasing resources
Fc	Fixed Pr resource installation or configuration cost
A	Option for configuration of Resources
N	Number of Cloudlets required
I	List of resource providers

$$DC = \sum_{i\epsilon I} Pr_i + \sum_{k=1}^{K} C_k + E \quad (5.1)$$

$$Pr_i = a \sum_{i=1}^{I} fr_i + fc_i \quad (5.2)$$

$$C_k = \sum_{k=1}^{K} n_k * C \quad (5.3)$$

$k = \{1, 2,..., K\}$ *number of V M s*

$a \in (0, 1)$ *Additional setup cost included by ISP*

$I = \{P\,r_1, P\,r_2,.., P\,r_n\}$ *resource providers* (*ISP or Cloud resource provider*)

The deployment cost (*DC* as given in equation 5.1) is the total sum of the leased resources from the providers (*Pr*), cost of a Cloudlet (*C*) and equipment (*E*). The equation takes into account that the Cloudlet owner might lease resources from different providers and that a different number of Cloudlets might be required depending on the size of an SME, as extended by equation 5.2 and 5.3. Some providers might include installation cost for leasing their resources, as presented in equation 5.2. The installation cost is a once-off fee represented by variable a. If the value of ‘a’ is 0 it means no installation cost is included otherwise it must be added to the cost for leasing provider resources.

The *DC* is important for ensuring the successful deployment of Cloudlet technology in environments such as a coffee shop. It is processed by the Cloudlet owner as stated on the CBM negative entity (costs). In ensuring a suitable deployment cost that a coffee shop owner can afford and also to avoid deployment failures, the following must be taken into consideration: Who are your customers?; how will they be using the Cloudlet?; how long will they be connected to the Cloudlet?, and lastly, how many customers at a time will be using a Cloudlet?.

The questions help in composing an SLA that is suitable for customers and fair to stakeholders. As indicated on the SLA the Cloudlet Consumers can enjoy a certain maximum number of connectivity hours to the Cloudlet. This will enable the Cloudlet owner to know how much resource is needed and what package they should lease from ISP to cater for all initiated tasks at a given time, which will also add to improving the consumers' QoE when consuming Cloudlet resources. In summary, this helps to identify the envisaged workload to be handled by the Cloudlet in the deployed environment.

The calculation of the workload (Cloudlet utilisation) requires the inclusion of the average task execution time of a Cloudlet, as stated in the SLA. This can be accomplished by using a queuing model (M/M/1), assuming that the request arrival rate follows a Poison Distribution (Adan, and Resing, 2001) depicted in equation p which is the product of the arrival rate (λ) with seconds (s) it takes a Cloudlet to process a a request. M/M/1 is a queuing model where the packet arrivals follow a Poisson process, the service times are exponentially distributed and there is one server. For an example, given that the average Cloudlet execution rate is 10 milliseconds (ms) and in an interval of 30 minutes during the peak period, approximately 150 requests are sent to a Cloudlet for processing. To calculate the rate of Cloudlet utilisation requires:

Changing ms to s: $10ms = 10 * 0.001 = 0.01$ *secs time it takes a Cloudlet to process one request*

Arrival rate of requests sent to a Cloudlet (using 150 tasks):

$$\lambda = \frac{150}{30 * 60} = 0.083 \textit{ requests/sec}$$

Therefore, the utilisation of a Cloudlet is:

$$p = \lambda * s = 0.083 * 0.01 = 0.00083$$

$$\textit{The utilisation rate lies in between}: 0 \leq p \leq 1$$

If the utilisation rate is equal to 1 it may require the Cloudlet owner to consider adding more resources such as storage and processing power to cater for the workload. Another cost that must be considered is the cost of equipment to enable consumers to have access to Cloudlet resources such as the wireless router, modem, and other equipment. However, for a business which already has a Wi-Fi service on their premises, they can reuse the equipment to support a Cloudlet deployment for Wi-Fi signal extending.

5.4 Comparing Costs and Benefits

The comparison between cost and benefits determines the feasibility of deploying a CBM in SMEs such as a coffee shop. It was conducted based on a Cloudlet owner because they are responsible for deploying a CBM in SMEs. The comparison was achieved using a Net Present Value (*NPV*) and a Cost-Benefit Ratio (*CBR*) to determine the feasibility of deployment. The *NPV* function is used to calculate the present value of an investment by the discounted sum of all cash flows received from a project (Taylor, 2015, Quah, and Haldane, 2007). The *NPV* is used to determine the *CBR* by dividing the *NPV* total with the deployment cost or investment cost, as shown in equation 5.5.

$$NP\ V = value\ (1 + r)^t \tag{5.4}$$

$$CBA\ Ratio = \frac{NP\ V_T otal}{DC} \tag{5.5}$$

An envisaged CBM feasibility using a coffee shop use-case considered the following information for an example: A coffee shop owner spends a total of R20 000 (including installation cost) in deploying Cloudlet technology on their premises. Based on the total costs spent on deployment they expect to increase the annual revenue by R 10 000 for the next three years. If the rate of inflation was 4.4%, each value is discounted as follows (using equation 5.4 and 5.5):

$$NP\ V\,(10000, year1) = 10000(1 + 0.044)1 = 10440$$

$$NP\ V\,(10000, year2) = 10000(1 + 0.044)2 = 10899.36$$

$$NP\ V\,(10000, year3) = 10000(1 + 0.044)3 = 11378.94$$

$$NP\,V_{Total} = 10440 + 10899.36 + 11378.94 = 32718.29$$

Calculating the *CBA ratio* using equation 4.5:

$$CBA\ Ratio = \frac{32718.29}{20000} = 1.64$$

Algorithm 1 An algorithm for interpreting the value of CBA Ratio

```
Input: CBA Ratio
Output: SMEs deployment decision
if CBARatio < 1.0 then
    Cost exceeds the benefits. A Cloudlet must not be deployed.
else if CBA Ratio == 1.0 then
    Cost equals to the benefits. A Cloudlet can be deployed but with little viability.
else
    Benefits exceeds the costs. A Cloudlet must be deployed.
endif
```

Algorithm1 is used in this work to interpret the results obtained from the calculated *CBA ratio,* which is 1.64 %. The 1.64% *CBA ratio* means that a coffee shop will make about 164 cents for every rand they spend on the deployed Cloudlet (operational cost) on their premises. Therefore, based on the above calculation the deployment of CBM in a coffee shop is feasible.

5.5 Chapter Summary

This chapter carried out a feasibility evaluation for deploying CBM using a coffee shop scenario. This chapter contributes to objective three of the study by proving the feasibility of the designed CBM. The feasibility was explored using *CBA*, the concepts of *NPV* and *CBR* were used to evaluate the feasibility of a Cloudlet in a coffee shop scenario. Two main topics were discussed in this chapter namely, the identification of costs and benefits and the comparison of costs and benefits. The difference between the above-mentioned two subtopics is that the first one looks at all role-players and the last one looks at the Cloudlet owner as the policymaker with the power to influence the deployment of a Cloudlet. The evaluation carried out proved that deploying a CBM in a small business environment such as a coffee shop is feasible. Therefore, Cloudlet deployment can be used as a mechanism to achieve financial stability and adaptation to competition levels. The following chapter presents the significance of the SLA concept as a means to mitigating Cloudlet deployment failures. The next Chapter shows how the use of an SLA will optimise operational costs of a CBM deployed in a coffee shop scenario.

Chapter 6

Service Level Agreement Scenario-Based Simulation

6.1 Introduction

In Chapter 4, the development process of an SLA-driven CBM and a use-case where this model can be applicable were covered. As mentioned in the previous chapters, SMEs cannot afford to make irresponsible financial decisions. This study considers both sharing of browser history data and the use of an SLA to minimise both the deployment and operational costs of a CBM for SMEs. This chapter demonstrates how SMEs can adapt to a CBM's operational costs through the use of an SLA to control the provisioning and allocation of resources in a network, thus eliminating service violations. This can further ensure service guarantees and improved QoE to Cloudlet consumers.

6.2 Basic Assumptions

The CloudSim treats the Cloudlets as mimics of Cloud tasks whereby, the host of the VM is treated as a Cloudlet instance running multiple VMs used for task processing. The tasks are distributed to different VMs on the network based on their size. In this simulation, 150 is an approximate number of tasks assumed to be initiated for processing during high peak hours in a Coffee Shop environment.

Since the tasks processed by the Cloudlet are tasks whose initiator has agreed to the terms (SLA) stated on the disclaimer page. The Cloudlet resources are consumed only when the SLA is accepted.

The violations of an SLA are caused by a shortage of resources (add more VMs) or by a network outage due to depletion of the Internet service provided by an ISP. In a case where some tasks are not processed, it is assumed that a Cloudlet owner receives a detailed notification of the violation, consisting of the percentage of met SLOs and possible issues that need to be attended to ensure service guarantees are met.

6.3 Description of the Simulation Setup

The simulation was carried out using the provided machine specifications and software listed in Table 6.1, running on an ASUS laptop. The CloudSim Plus tool was pulled from the GitHub. The GitHub repository is an open source online project repository used by engineers and developers to store and update code in (github.com/manoelcampos/cloudsim-plus). In this work, the CloudSim tool was imported to NetBeans IDE, configured with a maven package used to transform code into a JAR file that can be used to perform the tests. CloudSim Plus is an open source tool that is used for simulating Cloud scenarios through providing basic classes that describe data centres, virtual machines, applications, users, computational resources and policies for management (Silva Filho et al., 2017). The CloudSim tool was selected due to its functionality to enable the simulation of policies on a network to manage resource usage and allocation. This means that policies for management (SLA) classes were considered for this simulation.

The downloaded CloudSim Plus tool package using CloudSim 3 was imported as a maven project on NetBeans and built with dependencies in order to download all the packages required for the project to execute. The CloudSim Plus simulation tool already has classes that take into consideration the SLA implementation however, they require an edit to suit the coffee shop use-case scenario. Lastly, to check if the setup of the CloudSim Plus tool went successfully required the testing and compiling of files under the folder cloudsim-plus-examples. The files executed successfully.

6.3.1 SLA metrics and their dimensions

As mentioned in Chapter Three, an SLA is comprised of two parameters namely, the SLO and SLI. The SLA percentage in this work is described as an amount of SLO metrics the Cloudlet owner was able to meet, which are monitored by the use of SLI. The discussion in the subsections are the metrics considered for an SLA negotiation between the Cloudlet owner and the coffee shop customers (Cloudlet consumers) that determines the QoS based on the Cloudlet's performance:

6.3.1.1 Task Completion time

The task completion time refers to the time it took the VMs to process all the tasks pushed to them and it is measured in Milliseconds (ms). The threshold of the task completion time is stated as 100 ms, this refers to the envisaged acceptable latency when using a Cloudlet.

6.3.1.2 CPU utilisation

The CPU utilization is a percentage of compute power used to tasks in the VMs. The threshold of CPU utilisation is 99%. The higher the CPU utilization, the greater the % of SLA violation as this will degrade performance in the network.

6.3.1.3 Availability

Availability is the measure of resource accessibility on the network to cater for the processing and execution of tasks on the network. The threshold of resource availability is 100% starting from a minimum of 90%.

6.3.1.4 Wait time

The time difference between the start time and task completion time. Although the wait time is part of the SLA metrics, due to the use of a shared-time scheduling algorithm, the wait time was not considered during the simulation.

6.3.1.5 Throughput

Throughput is represented by the total number of tasks processed successfully by the VMs on the network.

6.3.1.6 Fault tolerance level

It is the measure of performance in terms of resource utilisation on the network. The threshold is set to none but the minimum is 2.

6.3.2 The main tasks carried out before the simulation

The simulation required the creation of an SLA contract that would be verified against values obtained as simulation performance metrics. The SLA contract was written in JSON format, it consisted of SLA metrics and their constants measured in different units as detailed above.

The classes namely, **CloudletTaskCompletionTimeWorkLoadMinimisationRunner** (subclass) and **CloudletTaskCompletionTimeWorkLoadMinimisationExperiment** (superclass) play an important role in the simulation of the SLA carried out in this work:

The CloudletTaskCompletionTimeWorkLoadMinimisationRunner class is responsible for distributing the tasks fairly between multiple VMs on the network. As mentioned in the assumptions, during peak hours an approximate number of 150 requests might be flooded to the Cloudlet. Taking into consideration that the weight of requests initiated might not be the same, some might relate to email account access while others might be for video calls. To take care of the above-mentioned consideration requires a mechanism to distribute the tasks fairly between multiple VMs on the network.

The table 6.1 lists the configuration properties and software of the used machine (ASUS laptop) for simulation.

Table 6.1: ASUS laptop system properties and software considered

System CPU	Intel (R) Core(TM) i7-4500U CPU @ 1.80GHz 2.40GHz
System RAM	8 GB
NetBeans Software version	8.2
Maven package Version	3.5.4
System Operating Sys-Tem	Windows 10 64-bit
Java Development Kit (JDK)	1.8.0

This is already taken care of by VM scheduling algorithms, the range of *10- 40 number of VMs (increasing with 10 units)* was used to observe how the SLA behaves, while the number of Cloudlet tasks and VM properties remained constant. This class is a subclass of CloudletTaskCompletionTimeWorkLoadMinimisationExperiment Class (as shown in Fig.6.2). It inherits properties of a superclass to achieve a successful task distribution and processing amongst VMs.

The CloudletTaskCompletionTimeWorkLoadMinimisationExperiment Class is a superclass of the runner class. The code skeleton shown in Fig. 6.1 was added to the CloudletMeetingCompletionTime method to raise flags when the SLA is violated, provide possible assumptions in a case where the SLA percentage is less than 100, and to calculate the session extension for unsuccessful tasks returns a double value shown in table 6.2.

```
double x=(150*cloudletCompletionTime.getMean())/totalOfcloudletSlaSatisfied;
double sessionExt=x-cloudletCompletionTime.getMean();
final double percent = totalOfcloudletSlaSatisfied * 100 / broker.getCloudletFinishedList().size();
System.out.printf("\n # Total of cloudlets satisfied SLA completion time of %.2f secs: %.0f of %d",
    getMaxTaskCompletionTime(broker), totalOfcloudletSlaSatisfied, broker.getCloudletFinishedList().size());
System.out.printf("\n # Percentage of cloudlets that complied with the SLA Agreement:  %.2f %%", percent);
if (percent<=90){
    System.out.printf("\n # The SLA Agreement was violated");
    System.out.printf("\n # Penalty: The session should be extended with "+ df1.format(sessionExt)+" ms for all unsuccessful tasks.");
    System.out.println("\n # The violations are due to resource (memory and CPU) shortage, add more resources");
}else if(percent>90 & percent<=94){
     System.out.printf("\n # The SLA Agreement was violated");
     System.out.printf("\n # Penalty: The session should be extended with "+ df1.format(sessionExt)+" ms for all unsuccessful tasks.");
     System.out.println("\n # The violations are due to resource (memory and CPU) shortage, add more resources");
}else if(percent>=95 & percent<=99){
    System.out.printf("\n # The SLA Agreement was successful");
    System.out.printf("\n # However, some tasks were not completed. Penalty: The session should be extended with "+ df1.format(sessionExt)+" ms for all uns
    System.out.println("\n # The violations are due to possible network outage, address the issue");
}else{
      System.out.printf("\n # The SLA Agreement was successful with no violations");

}
```

Figure 6.1: A function for calculating the percentage of tasks that met a completion time

```
public class CloudletTaskCompletionTimeWorkLoadMinimizationRunner extends ExperimentRunner<CloudletTaskCompletionTimeWorkLoadMi
    static final int[] VM_PES = {2, 4, 6};
    static final int[] VM_MIPS = {10000, 15000, 28000};
    public static final int VMS = 30;
    public static final int CLOUDLETS = 150;
```

Figure 6.2: A view of a CloudletTaskCompletionTimeWorkLoadMinimisation- Runner Class

The information in table 6.2 shows the results obtained during the simulation as the number of VMs were increased in a network to test the amount of SLOs that can be achieved.

Table 6.2: Simulation Results

No. of VMs	SLO %	No. of Tasks Completed	Throughput (%)	Obtained task completion time (ms)	Expected Completion time(ms)	Duration Extension (ms)
10	56.67	85	56.67	87.27	154.01	66.74
15	92	138	92	50.36	54.74	4.38
20	96.67	145	96.67	40.05	41.43	1.38
25	98	147	98	31.6	32.24	1.38
30	98	147	98	26.88	27.43	0.55
35	98	147	98	24.27	24.77	0.50
40	98	147	98	21.82	22.27	0.45
45	98	147	98	20.19	20.60	0.41
50	100	150	100	18.25	18.25	0.00

6.4 Results discussion

The results obtained from the simulation are presented in Fig. 6.3 - 6.9 and summarised by table 6.2. The first two figures (Fig. 6.3 and Fig. 6.4) show the effect of increasing the number of VMs in the network in alignment with the number of completed tasks and the number of Service Level Objectives (SLOs) achieved when VMs increase. The SLOs in this context refers to the agreed thresholds of QoS stated in the SLA. On the x-axis for both graphs, the number of VMs are presented. On the y-axis for Fig. 6.3 is the percentage of achieved SLOs and on Fig. 6.4 is the number of completed tasks by the VMs.

The figures (Fig. 6.5 to 6.9) shows how time relates to either the VMs increase or the SLA percentage achieved in the network. Fig.6.5 shows the variation between the achieved and the expected completion time of the task. Then Fig. 6.6 further details the relationship between the achieved task completion time and SLA percentage. Lastly, depending on the expected task completion time, Fig. 6.9 shows how the increase in resources eliminates the session extension time resulting from the expected completion time of the tasks.

6.5 Results Discussion

This section provides an interpretation of the results obtained from the simulation of an SLA. The simulation showed a directly proportional relationship between the percentage of an SLA status and the number of available resources in a network. This implies that as more resources were added to a network so there was an increase in the SLA status.

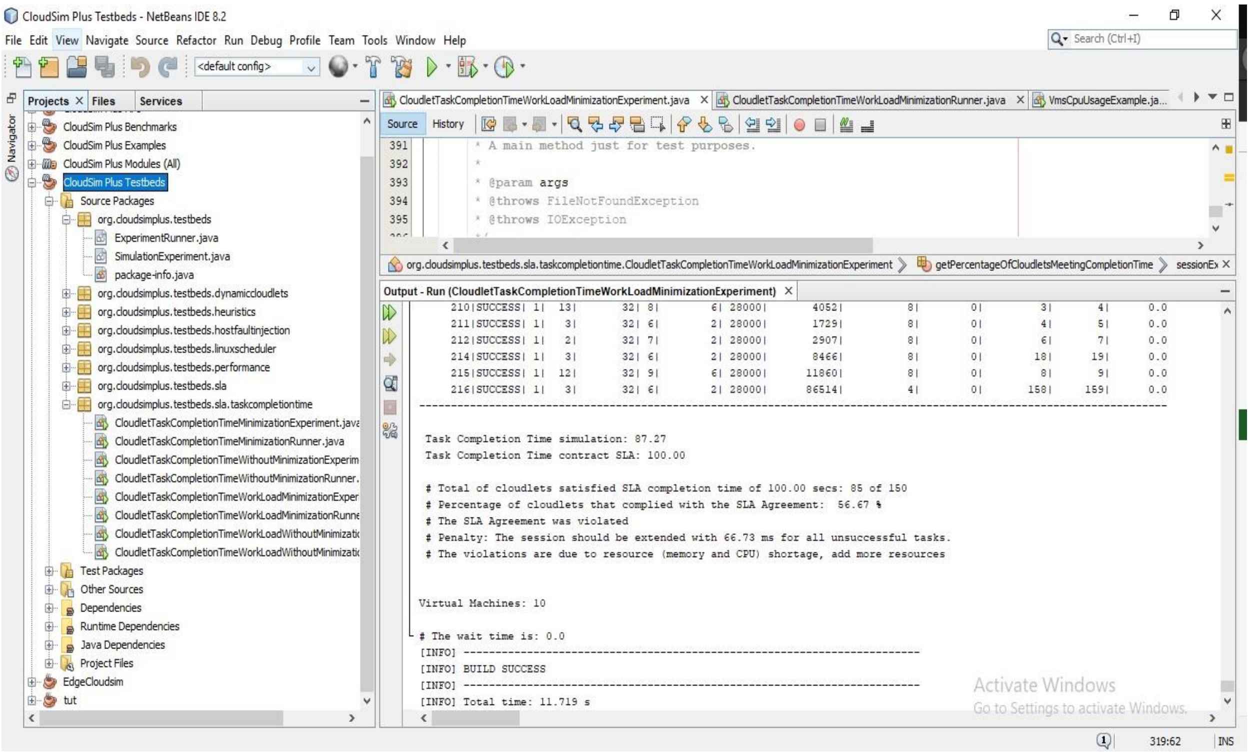

Figure 6.3: A screenshot of the simulation results of 10 VMs on a network

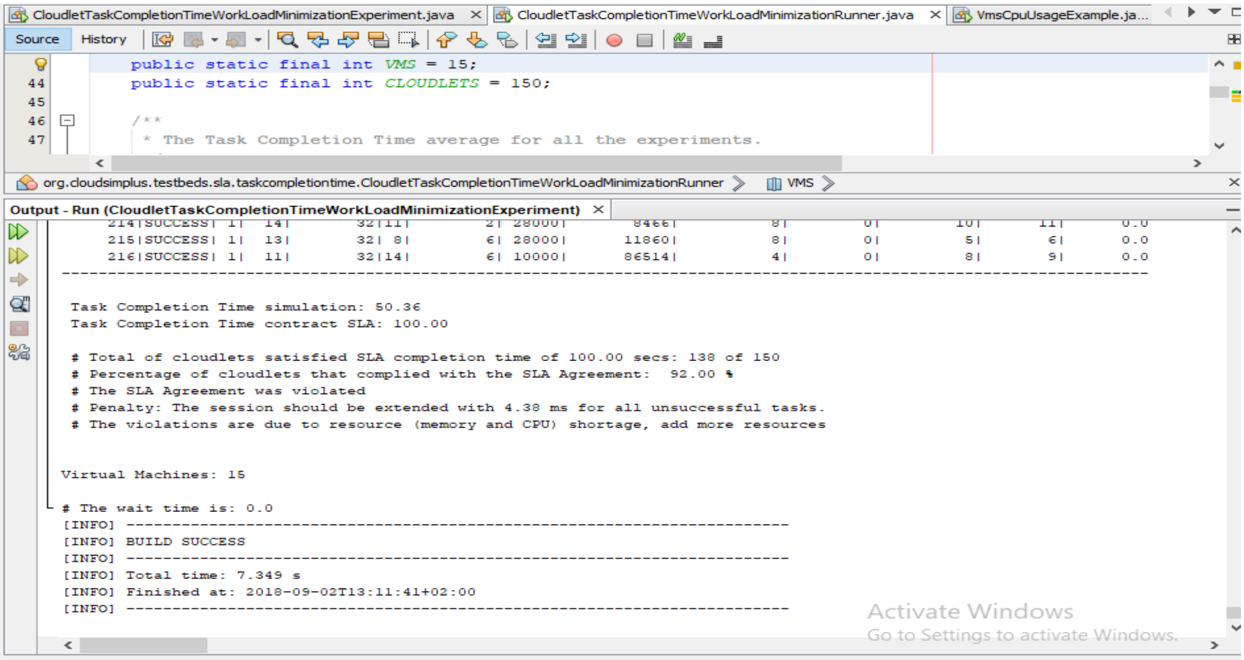

Figure 6.4: A screenshot of the simulation results of 15 VMs on a network

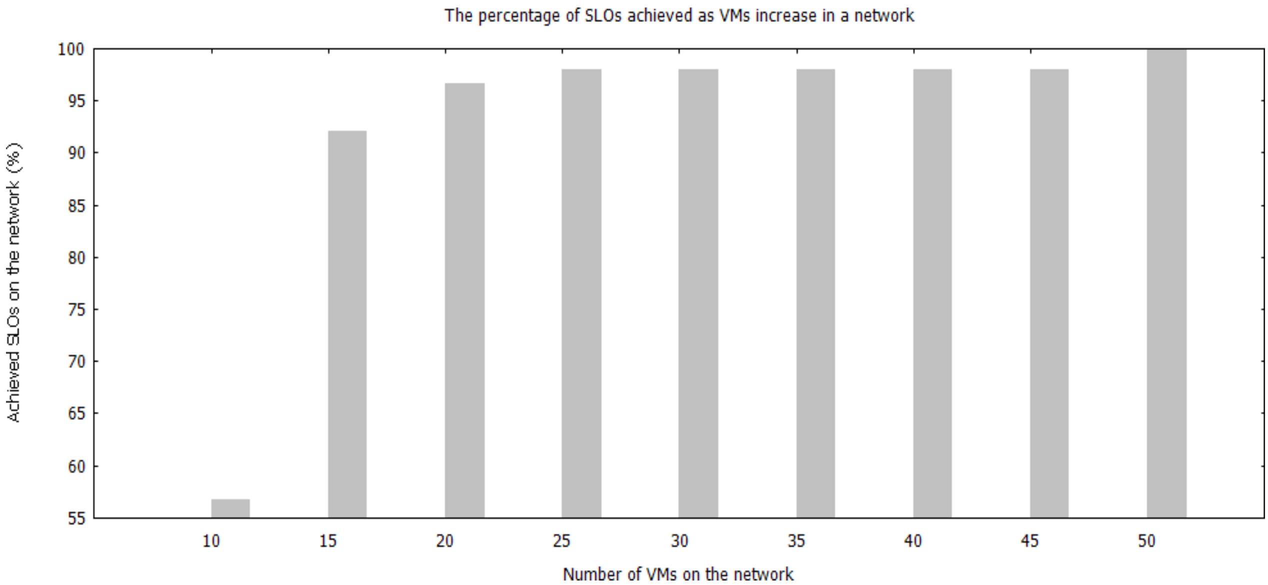

Figure 6.5: The effect of an SLA status as the VMs increases on the network

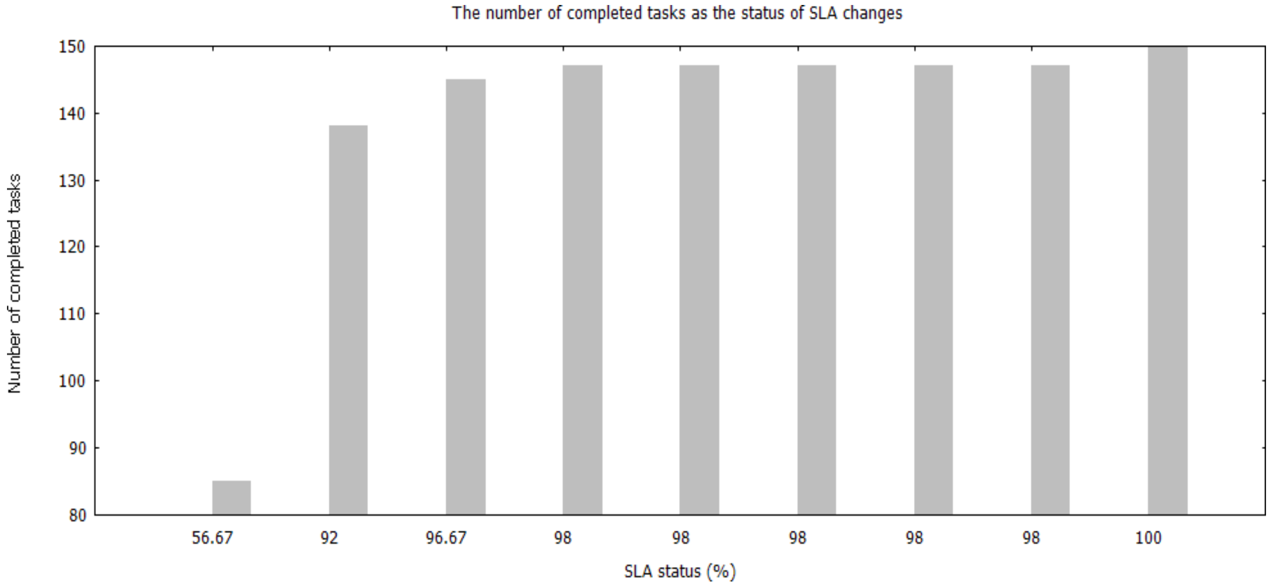

Figure 6.6: The relationship between completed tasks and the SLA status

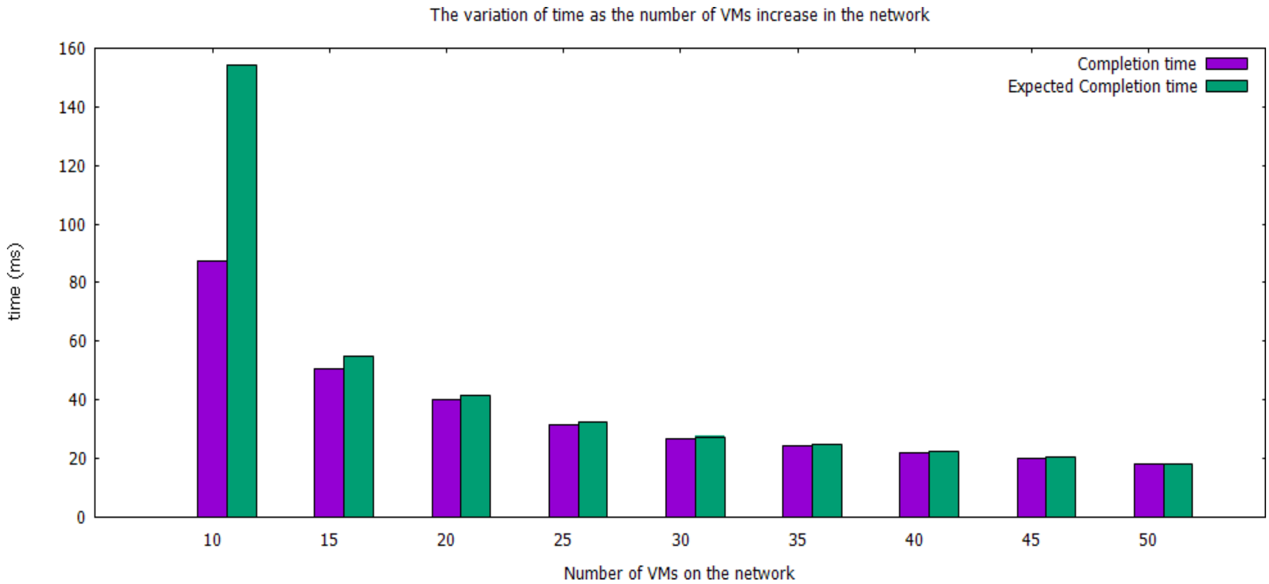

Figure 6.7: The relationship between completion time and expected time of tasks as VMs increase on the network

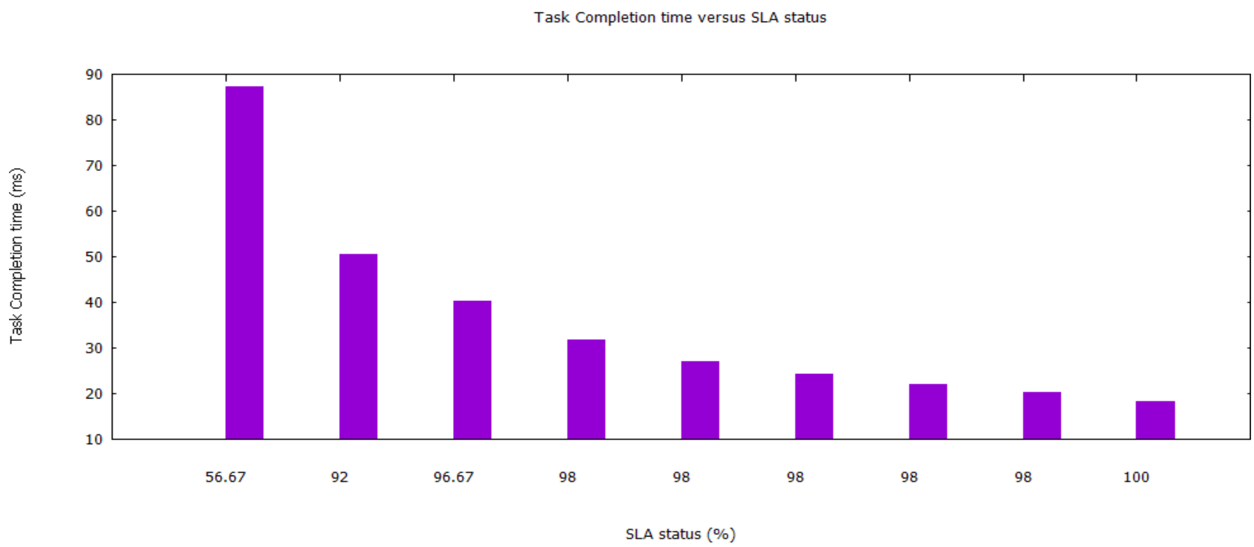

Figure 6.8: The relationship between task completion time and the SLA status

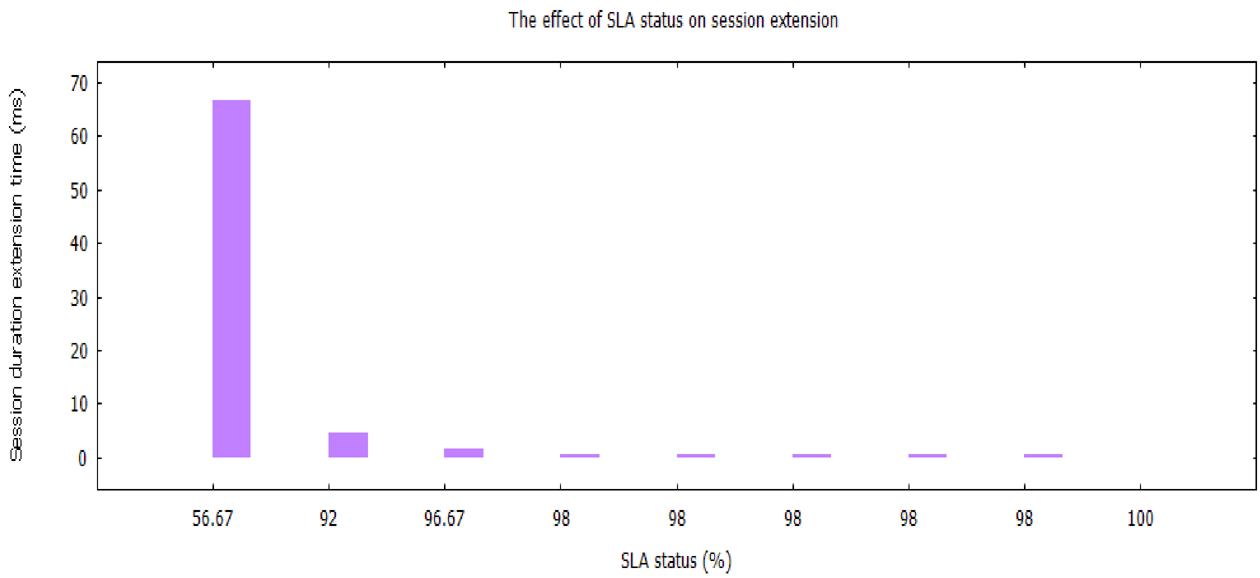

Figure 6.9: The relationship between service duration time and the SLA status

The figures Fig. 6.3 and Fig. 6.4 serve as evidence of the obtained simulation results shown in Netbeans IDE. For example Fig. 6.3 shows the results for a case where there are only 10 VMs in the network processing 150 tasks.

The simulation was repeated three times to ensure the quality of results and in all instances, the results remained the same. As mentioned in the simulation assumptions, the number of tasks processed by the VMs were kept constant *(150 tasks)*, only the number of VMs to process the tasks were increased. Fig. 6.5 shows that as the number of VMs were increased in a network, the percentage of the number of SLOs achieved also increased. The percentage of achieved SLOs is obtained through the comparison between obtained QoS simulation values of metrics such as task completion time and resource availability with that declared on the SLA contract. The comparison is controlled by the threshold values on the SLA QoS metrics. It was noticed that both graphs Fig. 6.5 and Fig. 6.6 possess a similar behavioural pattern.

This implies that as the number of SLOs achieved increased with an increase in VMs So too did the throughput value increase as shown in Fig. 6.6. The increase in both throughput and the SLA status eliminates SLA violations, ensures consumers' QoE and further minimises the operational costs of a Cloudlet.

In the event where the obtained SLA QoS metric values are higher than the thresholds stated on the SLA contract, the SLA status showed the effects (as shown in both Fig. 6.1 and Fig. 6.6). The effect of a single or more SLA metrics affects the number of tasks processed completely by the VMs, as shown in Fig. 6.6. For example, it was observed that although in many instances the completion task time was met but, due to other metrics not met, the SLA percentage decreased (shown in table 6.2 and Fig. 6.7). Although the number of VMs were continuously increased in the network both the throughput and SLA status belonging to VMs ranging from 25 to 45 remained the same. However, as shown in Fig. 6.7 the task completion time differed. The possible conclusion that can be drawn from this effect is that the tool limitations or network outage might have been the cause of this effect due to the fact that resources were increased but results remained the same for that number of VMs.

One of the most important factors in the SLA negotiations is the addressing of penalties due to failure in meeting the agreed SLOs on the SLA. The above is referred to as compensation for SLA violations. Few works have been published that focus on addressing compensation of SLA violations. One of these works includes a Cloud-based SLA availability framework that evaluates the SLAs of different Cloud Providers to help them achieve an effective penalty degree (Xia et al., 2013). The framework enables the calculation of both resource availability and penalty cost to determine a penalty degree that will help Cloud Providers achieve more profit. However, this work focuses on addressing only the violations made by consumers, not providers, which does not improve consumer QoE.

On the CBM, to improve QoE for consumers in cases as shown in Fig. 6.5 where less than 150 tasks are completed, a session duration is extended. The extension of session duration was calculated using equation 6.2. The equation 6.2 is the difference between the expected task completion time (*ECT*) and the task completion time (*CT*) obtained from the simulation results. Furthermore, the *ECT* was calculated using the equation 6.1, where the product of the total number of tasks pushed to VMs (*n*) and *CT* is divided by the number of tasks completed (*Cn*). The session extension serves as compensation to Cloudlet consumers for the unprocessed tasks. Although the above action will ensure QoE to Cloudlet consumers, it will increase the operational costs of the deployed CBM. Fig. 6.9 shows the session extension extended for all uncompleted tasks, the increase in session extension increases the operational cost of a Cloudlet. To minimise the operational costs as mentioned, the assumption is that the Cloudlet owner will receive a notification of SLA violations and the issues to be addressed (Suneel, and Guruprasad, 2015). Through attending to the notification, a coffee shop owner would be able to take countermeasures to help avoid SLA violations and also improve both Cloudlet performance and minimise its operational cost.

$$ECT = (n * CT)/Cn \tag{6.1}$$

$$SET = ECT - CT \tag{6.2}$$

The graph in Fig. 6.7 shows the variation between the *ECT* and the obtained *CT* as the number of VMs increased. As the number of VMs were increased in a network, it was observed that the completion and expected completion time slowly became equal which is a good thing when ensuring QoE for Cloudlet Consumers. This can be interpreted as to ensure that less session extension time (*SET*) is issued, the resources in place must be sufficient for all tasks requested for processing (as seen by the effect shown in fig. 6.9). By doing so, a Cloudlet owner will be able to save the cost of leasing resources while attracting more customers to the coffee shop.

6.6 Chapter Summary

The results provided in this chapter shows that the number of resources available when requests are initiated during peak hours have an effect on the success of an SLA. These results demonstrate the effect of an SLA in avoiding service violations through monitoring the provisioning and allocation of resources on a network. Through the above measures, SMEs can be able to avoid SLA violations, in short, minimising CBM operational costs. Also outlined are possible issues that can give rise to SLA violations namely, network outage or lack of resources (memory and processing power). Being aware of the root of SLA violations can help SMEs mitigate violations and operational costs. The conclusion can be drawn that through the deployment of SLA-based CBM in SMEs such as coffee shop can help the business policy-makers take informed decisions in ensuring service guarantees to Cloudlet consumers are met, and also improve both marketing and purchasing patterns of their core product, due to the deployment of a low latency, resource-rich, high bandwidth, one-hop Internet connection that is guided by an SLA.

Chapter 7

Conclusion and Future Work

7.1 Introduction

This research was conducted to address the gap in the literature with regards to the lack of CBM that uses an appropriate price function to negotiate the satisfaction of all role-players in the business model. The lack of CBM in literature makes it difficult for businesses categorised as SMEs to trust that the deployment of the Cloudlet can yield benefits for them and value to their customers. Since SMEs are prone to financial problems, to deploy or adopt a new service on their premises requires the evaluation of the service's feasibility. In this study, a CBM to ensure stakeholders' service guarantees were met was developed.

The development of a CBM involved an intensive investigation of strategies that could be used to ensure fair negotiations amongst involved role-players/ stakeholders namely, the Cloudlet consumer, Cloudlet owner, ISP or Cloud Provider. The strategy is called an SLA. In this study, the SLA is used to monitor the Cloudlet service provision to ensure that at all times the services are being guaranteed. Furthermore, the composition of a CBM's components took into consideration the capturing of users' browser history, its analysis, and management. The shared browser history data by Cloudlet consumers (fixed non-monetary value) is referred to as an appropriate price function that attracts SMEs customers (due to the provision of QoE by the deployed Cloudlet), improves resource efficiency in the network and reduces energy costs (operational cost) for Cloudlet owners.

7.2 Summary/ conclusion

The feasibility study based on the coffee shop scenario using the cost-benefit ratio gave a 2.93% result. Which means that a coffee shop can make about 293 cents for every rand they spend on the operations of the deployed Cloudlet. The results provide enough evidence to state that the deployment of a CBM in SMEs such as a coffee shop will yield more benefits and service guarantees as stated on the SLA between involved role-players in the model. Also, the SLA results provided in Chapter Five gave informative knowledge on how the Cloudlet owner can ensure service guarantees (by managing resources in the network) to Cloudlet consumers and at the same time avoid high service violations that can lead to deployment failure.

Therefore, it can be concluded using the above two results that the adoption of a Cloudlet Business Model by Small, Medium Enterprises can help SMEs to gain financial stability, enable adjustment to competition levels amongst SMEs, and promote good marketing of the business and core product. To consumers, the adoption can bring about value because it enables them to access customised online services due to browser history data sharing. This implies that the Cloudlet consumers will experience high QoE when accessing Cloud resources using a one-hop connection without having to worry about high latency. The possibility of the two benefit perspectives is due to the deployment of a low latency, context-aware, non-monetary Wi-Fi connection.

Subsection 1.6.2 of Chapter One presented four objectives designed to achieve the main goal of the study. The below discussion details how the four objectives were achieved.

OBJECTIVE I: To investigate existing SLA-based Cloud business models or similar models that can be aligned with a Cloudlet environment considering appropriate price functions.

This was achieved through an intensive investigation in the literature to find existing models that use an SLA to ensure management and provisioning of resource in both Cloud and Grid environments. The models found in literature provided guidelines to assist in developing an SLA-based CBM. Part of the investigation was to look at existing methods used to bill resource allocation or provision of service in the Cloudlet environment, which is useful for SLA negotiations between Cloudlet owners and the resource providers. It was ascertained that due to the lack of resources as compared to the Cloud, the Cloudlet uses competitive based price function such as auctioning and supply-demand. This knowledge was very useful since this research considers the use of two price functions to ensure stakeholders service guarantees. It was observed that the supply-demand price function can be used between the Cloudlet owner and Cloud Provider or ISP to ensure the leasing of efficient resources that will ensure service guarantees later improve the customers' QoE and SMEs QoB. However, the supply-demand is not applicable to the negotiations between the Cloudlet consumer and the Cloudlet owner. A fixed non-monetary value referred to as the sharing of users' browser history is used for negotiations between the Cloudlet consumer and Cloudlet owner.

OBJECTIVE II: To explore how the use of small data sensors can be beneficial for an SLA-based Cloudlet Business Model.

The main purpose of this objective was to enable the discovery of existing mechanism used to capture, analyse and manage context. Firstly, the definition of context awareness as based on this study was considered. Context-awareness in this study refers to the ability of the Cloudlet to use user browser history data to provide users with customised online services. It was obtained in the literature that although the compensation of Cloud-routers in SMEs is free, the SMEs' owners tend to increase the core product prices in order to maintain the deployment of the Cloud-routers. This course of action is not friendly to SMEs' customers. Through investigation, it was observed that the gathering of small data can be used to help reduce the cost of maintaining the Cloudlet deployment. Because a context-aware Cloudlet minimises task processing by prefetching data from the cache to the task initiator, minimising task processing reduces the amount of power a Cloudlet use to process each task further, reducing the cost of maintaining the Cloudlet deployment. This develops a CBM into an attractive business model.

OBJECTIVE III: To design a proposed Cloudlet Business Model that integrates SLAs of role-players with appropriate pricing functions.

The design of the CBM took into consideration the information and knowledge gained from both Objectives I and II. The afore-mentioned objectives enabled the composition of the model components, their processes, and connection to ensure stakeholders service guarantees. The SLA plays a significant role in the CBM, it is used to capture the required roles of the role-players to service guarantees. Two SLA instances (SLA1 and SLA2) are considered in this model. SLA1 captures the responsibilities of the Cloudlet owner and Cloudlet consumer and uses a fixed non-monetary price function (pay less for bulk usage). SLA2 is used between the Cloudlet owner and resource providers such as an ISP or Cloud Provider and it uses a supply-demand price function to capture the responsibilities of role-players phrased as SLOs in the SLA. The difference between SLA1 and SLA2 is that SLA2 is for resource-leasing which can improve QoS and QoB but SLA1 is for Cloudlet consumption, which intends to determine the consumers QoE and the free will to share their browser history data. The integration of the two SLAs is done to ensure service guarantees during the deployment of a Cloudlet by SMEs. Furthermore, a scenario-based SLA simulation using SLA1 was conducted and the results showed that through the monitoring of service provision SMEs can eliminate service violations as well as improve both QoE and QoB.

A CBA tool was used to evaluate the feasibility of deploying a CBM in a coffee shop environment. Part of the evaluation involved the discovery of Cost and benefits parameters as per the role-players in a Cloudlet Business Model. The CBA considers the policymakers to be responsible for the deployment of this model, in this case, the Cloudlet owner. To find out how much the coffee shop will make due to the deployment of a CBM, a CBA ratio was calculated. The CBA ratio considers the NPV calculated based on the period of 3 years. The results showed that deploying a CBM in a coffee shop can yield to 164 cents for every rand spent in maintaining the deployment. Therefore, the CBM is feasible.

OBJECTIVE IV: To evaluate the effect an SLA can impose on a CBM using a coffee shop scenario.

The simulation was carried out using a CloudSim Plus tool running on a Netbeans 8.0 IDE in a Windows 10 64 bit system. The aim of the simulation was to demonstrate the effect an SLA has on management and provision of resources in a network using a coffee shop scenario. The simulation took into consideration the following SLA QoS metrics: Task completion time, CPU utilisation, availability, throughput, and fault-tolerance level. The QoS metrics and their associated thresholds were defined and declared in a json file format called SLA.json. The final value of an SLA was determined based on the thresholds of QoS metrics and the network setup (no. of VMs and tasks). The simulation setup considered the range of 10-50 VMs while the tasks initiated were kept constant (250 tasks). The results obtained from the simulation indicated that as the number of VMs increased in the network, so too the SLA percentage increased. This implies that through using an SLA, a coffee shop can achieve an effective approach to the provision of Cloudlet resources by the edge without having to spend more money. The simulation also considered the issue of addressing penalties. Based on the SLA percentage obtained, a computation was designed to calculate session extension time to compensate edge users for violated services. By compensating Cloudlet consumers in cases of SLA violations a coffee shop can gain more Cloudlet consumers who would also want to consume the offered core services, example coffee.

7.3 Limitations and Future work

The lack of equipment to conduct a testbed of a Cloudlet instance in a coffee shop scenario limited this study from obtaining effective results in proving the feasibility of the model. Additionally, the lack of open source platforms with good documentation affected the study. Further work that can be conducted in this study is related to the security aspect of the model to ensure that the captured user browser history does not violate privacy-related issues. This can be done using encryption algorithms to ensure that the data captured is protected from sniffing and middleman attacks. The use of multi text label classification in machine learning can also be used to find out how SMEs can further increase benefits related to the captured user Browser history cached by the Cloudlet to provide customized services to SMEs. The idea is that the browser history data consists of features that can be extracted and used to determine and cluster consumers' interests to boost service provision and marketing.

References

Adan, I., and Resing, J. (2001), ‘Queuing theory’, *Eindhoven University of Technology* **3**, 1–4.

Ai, Y., Peng, M. and Zhang, K., 2018. Edge computing technologies for Internet of Things: a primer. Digital Communications and Networks, 4(2), pp.77-86.

Ala-Laurila, J., Mikkonen, J. and Rinnemaa, J., 2001. Wireless LAN access network architecture for mobile operators. IEEE Communications Magazine, 39(11), pp.82-89.

Alam, M. and Shakil, K.A., 2016. Big Data Analytics in Cloud environment using Hadoop. arXiv preprint arXiv:1610.04572.

Antonić, A., Marjanović, M., Pripužić, K. and Žarko, I.P., 2016. A mobile crowd sensing ecosystem enabled by CUPUS: Cloud-based publish/subscribe middleware for the Internet of Things. Future Generation Computer Systems, 56, pp.607-622.

Badidi, E., 2016. A broker-based framework for integrated sla-aware saas provisioning. arXiv preprint arXiv:1605.02432.

Belton, L.T., Beaty, B., Morris, T.H., Rodgers, D.S. and Smith, L.A., Wells Fargo Bank NA, 2015. Wireless multi-factor authentication with captive portals. U.S. Patent 9,088,891.

Bilal, K., Khalid, O., Erbad, A. and Khan, S.U., 2018. Potentials, trends, and prospects in edge technologies: Fog, cloudlet, mobile edge, and micro data centers. Computer Networks, 130, pp.94-120.

Bonomi, F., Milito, R., Natarajan, P. and Zhu, J., 2014. Fog computing: A platform for internet of things and analytics. In Big data and internet of things: A roadmap for smart environments (pp. 169-186). Springer, Cham.

Bonomi, F., Milito, R., Zhu, J. and Addepalli, S., 2012, August. Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing (pp. 13-16). ACM.

Borcoci, E., 2016, August. Fog Computing, Mobile Edge Computing, Cloudlets-which one. In SoftNet Conference.

Botta, A., De Donato, W., Persico, V. and Pescapé, A., 2016. Integration of cloud computing and internet of things: a survey. Future generation computer systems, 56, pp.684-700.

Buyya, R. and Bubendorfer, K. eds., 2010. Market-oriented grid and utility computing. John Wiley and Sons.

Cant, M.C. and Wiid, J.A., 2013. Establishing the challenges affecting South African SMEs. International Business and Economics Research Journal (IBER), 12(6), pp.707-716.

Casola, V., De Benedictis, A., Rak, M., Aversano, G. and Villano, U., 2013, December. An SLA-based Approach to Manage Sensor Networks as-a-Service. In 2013 IEEE 5th International Conference on Cloud Computing Technology and Science (Vol. 1, pp. 191-197). IEEE.

Christensen, C.M. and Johnson, M.W., 2009. What are Business Models, and how are They Built?. Harvard Business School.

Cloudera, 2014. Hadoop for small data. URL: https://vision.cloudera.com/hadoop-for-small-data/

Dar, S., Franklin, M.J., Jonsson, B.T., Srivastava, D. and Tan, M., 1996, September. Semantic data caching and replacement. In VLDB (Vol. 96, pp. 330-341).

del–Río–Ortega, A., Gutiérrez, A.M., Durán, A., Resinas, M. and Ruiz–Cortés, A., 2015, June. Modelling service level agreements for business process outsourcing services. In International Conference on Advanced Information Systems Engineering (pp. 485-500). Springer, Cham.

Dolui, K. and Datta, S.K., 2017, June. Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing. In 2017 Global Internet of Things Summit (GIoTS) (pp. 1-6). IEEE.

Dou, Z. and Heinzelman, W.B., 2013. Benefits of utilizing an edge server (cloudlet) in the mocha architecture (Doctoral dissertation, University of Rochester. Department of Electrical and Computer Engineering).

Duro, F.R., Blas, J.G., Higuero, D., Perez, O. and Carretero, J., 2015. CoSMiC: A hierarchical cloudlet-based storage architecture for mobile clouds. Simulation Modelling Practice and Theory, 50, pp.3-19.

E-Imports, E. B. S., 2017. Coffee statistics. URL: http://www.e-importz.com/coffee-statistics.php

Ewer, T., 2017. 6 of the Biggest Small Business Challenges (And How to Overcome Them). URL: https://www.hiveage.com/blog/challenges-facing-small-businesses/

Fielt, E., 2014. Conceptualising business models: Definitions, frameworks and classifications. Journal of Business Models, 1(1), pp.85-105.

Gao, Y., Hu, W., Ha, K., Amos, B., Pillai, P. and Satyanarayanan, M., 2015. Are cloudlets necessary?. School Comput. Sci., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-CS-15-139.

Gordon, S. (2014), 5 Ways Small Data Can Be More Valuable than Big Data. URL: https://www.destinationcrm.com/Articles/ReadArticle.aspx?ArticleID=98927

Guo, B., Wang, Z., Yu, Z., Wang, Y., Yen, N. Y., Huang, R. , and Zhou, X. (2015), 'Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm', ACM Computing Surveys (CSUR) 48(1), 7.

Herlocker, J.L., Konstan, J.A. and Riedl, J., 2000, December. Explaining collaborative filtering recommendations. In Proceedings of the 2000 ACM conference on Computer supported cooperative work (pp. 241-250). ACM.

Hiertz, G.R., Denteneer, D., Max, S., Taori, R., Cardona, J., Berlemann, L. and Walke, B., 2010. IEEE 802.11 s: the WLAN mesh standard. IEEE Wireless Communications, 17(1), pp.104-111.

Hou, Z., 2018. Efficient Data Access in Cloudlet-based Mobile Cloud Computing (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

Hu, W., Gao, Y., Ha, K., Wang, J., Amos, B., Chen, Z., Pillai, P. and Satyanarayanan, M., 2016, August. Quantifying the impact of edge computing on mobile applications. In Proceedings of the 7th ACM SIGOPS Asia-Pacific Workshop on Systems (p. 5). ACM.

Jararweh, Y., Ababneh, F., Khreishah, A. and Dosari, F., 2014. Scalable cloudlet-based mobile computing model. Procedia Computer Science, 34, pp.434-441.

Jenkins, G.P., 1999. Evaluation of stakeholder impacts in cost-benefit analysis. Impact Assessment and Project Appraisal, 17(2), pp.87-96.

Jin, A.L., Song, W., Wang, P., Niyato, D. and Ju, P., 2015. Auction mechanisms toward efficient resource sharing for cloudlets in mobile cloud computing. IEEE Transactions on Services Computing, 9(6), pp.895-909.

Kbar, G. and Mansoor, W., 2006. Distributed Resources Management in Wireless LANs. International Journal of Business Data Communications and Networking (IJBDCN), 2(4), pp.46-58.

Khattak, A. B., and Jalal, A., 2015. Performance based service level agreement in cloud computing. Research Journal of Science and IT Management 4(4), 20–30.

Kotevska, O., Lbath, A. and Bouzefrane, S., 2016. Toward a real-time framework in cloudlet-based architecture. Tsinghua Science and Technology, 21(1), pp.80-88.

Krueger, M.W. and Dix, M., 2014. Analysis of Cloud Business Models for the Industrial B2B Market. IT Convergence Practice (INPRA), 2(2).

Lindstrom, M., 2016. Small data: the tiny clues that uncover huge trends. St. Mar- tin's Press.

Liu, J., Shen, H., Narman, H.S., Chung, W. and Lin, Z., 2018. A survey of mobile crowdsensing techniques: A critical component for the internet of things. ACM Transactions on Cyber-Physical Systems, 2(3), p.18.

Loke, S.W., 2012. Supporting ubiquitous sensor-cloudlets and context-cloudlets: Programming compositions of context-aware systems for mobile users. Future Generation Computer Systems, 28(4), pp.619-632.

Nandimath, J., Banerjee, E., Patil, A., Kakade, P., Vaidya, S. and Chaturvedi, D., 2013, August. Big data analysis using Apache Hadoop. In 2013 IEEE 14th International Conference on Information Reuse, and Integration (IRI) (pp. 700-703). IEEE.

Mach, P. and Becvar, Z., 2017. Mobile edge computing: A survey on architecture and computation offloading. IEEE Communications Surveys, and Tutorials, 19(3), pp.1628-1656.

Marešová, P. and Půžová, K., 2014. Application of the Cost Benefit Analysis Method in Cloud Computing in the Czech Republic. Procedia-Social and Behavioral Sciences, 109, pp.674-678.

Maresova, P., Sobeslav, V. and Krejcar, O., 2017. Cost–benefit analysis–evaluation model of cloud computing deployment for use in companies. Applied Economics, 49(6), pp.521-533.

McGilvary, G.A., Rius, J., Goiri, Í., Solsona, F., Barker, A. and Atkinson, M., 2013, December. C2ms: Dynamic monitoring and management of cloud infrastructures. In 2013 IEEE 5th International Conference on Cloud Computing Technology and Science (Vol. 1, pp. 290-297). IEEE.

Middleton, S. E., Surridge, M., Nasser, B. I., and Yang, X., 2009. Bipartite electronic SLA as a Business Framework to Support Cross-organisation Load Management of real-time Online Applications, in European Conference on Parallel Processing, Springer, pp. 245–254.

Modh, R.M. and Patel, J.M., 2013. A study on improving challenges of network with cloudlets for mobile cloud computing. Journal of global research in computer science, 4(4), pp.143-146.

Mor, Y., 2014, 5 reasons you need to process small data with hadoop. URL: https://www.xplenty.com/blog/process-small-data-hadoop/

Musa, S.M., Yousif, A. and Bashi, M.B., 2016. Sla violation detection mechanism for cloud computing. International Journal of Computer Applications, 133(6), pp.8-11.

Neves, P.C., Schmerl, B.R., Cámara, J. and Bernardino, J., 2016, January. Big Data in Cloud Computing: Features and Issues. In IoTBD (pp. 307-314).

Padilla, R.S., Milton, S.K. and Johnson, L.W., 2015. Components of service value in business-to-business Cloud Computing. Journal of Cloud Computing, 4(1), p.15.

Pang, Z., Sun, L., Wang, Z., Tian, E. and Yang, S., 2015, November. A survey of cloudlet based mobile computing. In 2015 International Conference on Cloud Computing and Big Data (CCBD) (pp. 268-275). IEEE.

Panigrahi, C.R., Tiwary, M., Pati, B. and Misra, R., 2015, August. CAB: cloudlets as agents of cloud brokers. In 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 381-386). IEEE.

Patel, P., Ranabahu, A. H., and Sheth, A. P., 2009. Service level agreement in cloud computing'.

Podlipnig, S., and B¨osz¨ormenyi, L., 2003. A survey of web cache replacement strategies, ACM Computing Surveys (CSUR) 35(4), 374–398.

Purdy, K., 2011. 'Run your home network like a coffee shop for easier guest access and family-friendly browsing'.

URL: https://lifehacker.com/5791208/run-your-home-network-like-a-coffee-shop?tag=kids

Quah, E., and Haldane, J., 2007. Cost-benefit analysis, Routledge.

Raei, H. and Yazdani, N., 2017. Performability analysis of cloudlet in mobile cloud computing, Information Sciences 388, 99–117.

Rawadi, J., Artail, H. and Safa, H., 2014, April. Providing local cloud services to mobile devices with inter-cloudlet communication. In MELECON 2014-2014 17th IEEE Mediterranean Electrotechnical Conference (pp. 134-138). IEEE.

Satyanarayanan, M., 2015. A brief history of cloud offload: A personal journey from odyssey through cyber foraging to cloudlets, GetMobile: Mobile Computing and Communications 18(4), 19–23.

Satyanarayanan, M. (2017), Edge computing for situational awareness, in 'Local and Metropolitan Area Networks (LANMAN), 2017 IEEE International Symposium on', IEEE, pp. 1–6.

Satyanarayanan, M., Bahl, V., Caceres, R., and Davies, N., 2009. The case for vm- based cloudlets in mobile computing, IEEE pervasive Computing.

Satyanarayanan, M., Schuster, R., Ebling, M., Fettweis, G., Flinck, H., Joshi, K., and Sabnani, K., 2015. An open ecosystem for mobile-cloud convergence. IEEE Communications Magazine 53(3), 63–70.

Serrano, D., Bouchenak, S., Kouki, Y., de Oliveira Jr, F.A., Ledoux, T., Lejeune, J., Sopena, J., Arantes, L. and Sens, P., 2016. SLA guarantees for cloud services. Future Generation Computer Systems, 54, pp.233-246.

Shaukat, U., Ahmed, E., Anwar, Z., and Xia, F. (2016), 'Cloudlet deployment in local wireless networks: Motivation, architectures, applications, and open challenges', Journal of Network and Computer Applications 62, 18–40.

Silva Filho, M. C., Oliveira, R. L., Monteiro, C. C., In´acio, P. R., and Freire, M. M., 2017. Cloudsim plus: a cloud computing simulation framework pursuing software engineering principles for improved modularity, extensibility and correctness, in 'Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium on', IEEE, pp. 400–406.

Suneel, K., and Guruprasad, D. H., 2015. A novel approach for sla compliance monitoring in cloud computing, International Journal of Innovative Research in Advanced Engineering 2(2), 154–159.

Tang, B., Gupta, H., and Das, S. R., 2008. Benefit-based data caching in ad hoc networks, IEEE transactions on Mobile Computing 7(3), 289–304.

Taylor, C., 2015. How to determine whether the cost-benefit ratio is positive or negative. URL: https://smallbusiness.chron.com/determine-whether-costbenefit-ratio-positive-negative-48312.html

Tiwary, M., Sahoo, K. S., Sahoo, B., and Misra, R., 2017. Cps: a dynamic and distributed pricing policy in cyber foraging systems for fixed state cloudlets, Computing 99(5), 447–463.

Verbelen, T., Simoens, P., De Turck, F., and Dhoedt, B., 2012. Cloudlets: Bringing the cloud to the mobile user, in Proceedings of the third ACM workshop on Mobile cloud computing and services, ACM, pp. 29–36.

Wollenick, N., 2012. Identifying Challenges in Applying the Value Co-Creation Approach in Practice:-a Case Study in the B2B Service Context.

Wu, H., Peng, Y., Long, K., Cheng, S., and Ma, J., 2002. Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement, in 'INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE', Vol. 2, IEEE, pp. 599–607.

Wu, Y., and Ying, L., 2015. A cloudlet-based multi-lateral resource exchange frame- work for mobile users, in Computer Communications (INFOCOM), 2015 IEEE Conference on, IEEE, pp. 927–935.

Xia, Q., Liang, W., and Xu, W., 2013. Throughput maximisation for online request admissions in mobile cloudlets, in Local Computer Networks (LCN), 2013 IEEE 38th Conference on, IEEE, pp. 589–596.

Xiong, K., and Chen, X., 2015. Ensuring cloud service guarantees via service level agreement (sla)-based resource allocation, in Distributed Computing Systems Workshops (ICDCSW), 2015 IEEE 35th International Conference on, IEEE, pp. 35–41.

Yu, W., Liang, F., He, X., Hatcher, W. G., Lu, C., Lin, J., and Yang, X., 2018. A survey on the edge computing for the Internet of things, IEEE Access 6, 6900–6919.

Zhang, J., Xiong, T., and Lou, W., 2014. Community clinic: Economizing mobile cloud service cost via cloudlet group, in Mobile Ad Hoc and Sensor Systems (MASS), 2014

IEEE 11th International Conference on, IEEE, pp. 208–216.

Zhijun, H., Robson, E. , and Boukerche, A. (2017), Towards efficient data access in mobile cloud computing using pre-fetching and caching, in ‘Communications (ICC), 2017 IEEE International Conference on’, IEEE, pp. 1–6.

Appendix

Appendix A

Xiaomi Smart Router

It is a wireless router that is used as a lightweight edge server equipped with a Linux-based operating system and a storage up to 6TB. It consist of three features namely, web page prefetching, movie prefetching and file backup applications Pang et al., 2015.

1. Web page prefetching

The Xiaomi Smart Router has a Sougou prefetching engine which enables the prediction of user future access based on the historical access log obtained from the engine. The prefetching of the predicted content enables the acceleration of web browsing.

2. Movie Prefetching

The router enables users to indicate the video of interests and automatically downloads them for later watch.

3. File Backup

Enable users store file such as documents and images on the router. The user can access the files remotely via an Internet.