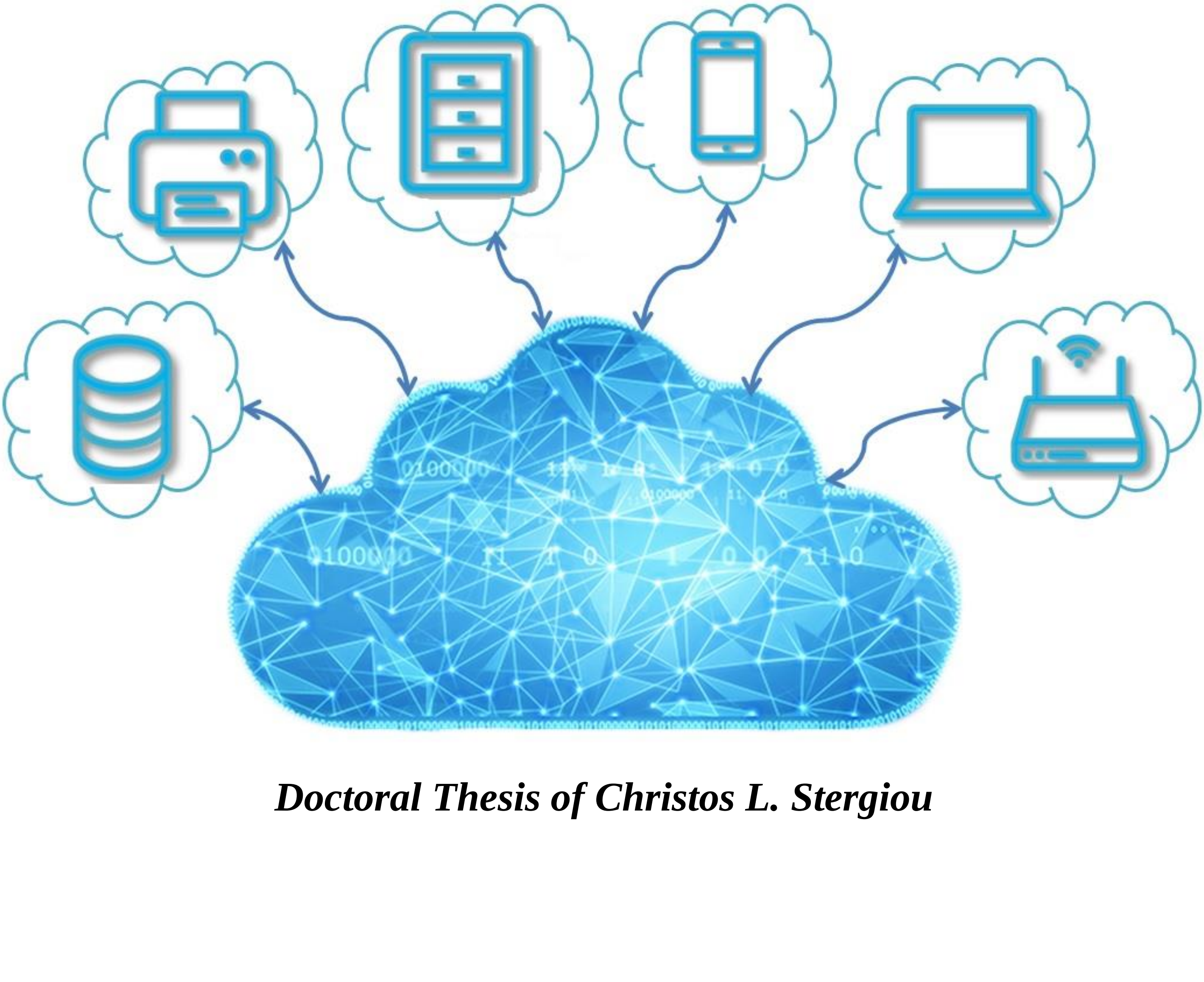
University of Macedonia
Department of Applied Informatics
Postgraduate Program Studies

Efficient and Secure Algorithms for Big Data Handling, Processing, and Delivery in Cloud Computing for Internet of Things Networks



Doctoral Thesis of Christos L. Stergiou

Thessaloniki, 2021

University of Macedonia
Department of Applied Informatics
Postgraduate Program Studies

Efficient and Secure Algorithms for Big Data Handling, Processing, and Delivery in Cloud Computing for Internet of Things Networks

Doctoral Thesis of Christos L. Stergiou

Doctoral Thesis Committee:
Konstantinos E. Psannis, Associate Professor, University of Macedonia (Advisor)
Georgios Evaggelidis, Professor, University of Macedonia
Theodoros Kaskalis, Associate Professor, University of Macedonia

Thessaloniki, 2021

“Efficient and Secure Algorithms for Big Data Handling, Processing, and Delivery in Cloud Computing for Internet of Things Networks”

Christos L. Stergiou

MSc, Wireless Communication Systems, Brunel University, 2012

Thesis submitted

in partial fulfillment of the requirements for the
Doctorate degree in Computer Science

Examination Committee:

____
Konstantinos E. Psannis, Associate Professor, University of Macedonia

____
Georgios Evaggelidis, Professor, University of Macedonia

____
Theodoros Kaskalis, Associate Professor, University of Macedonia

____
Konstantinos Margaritis, Professor, University of Macedonia

____
Manos Roumeliotis, Professor, University of Macedonia

____
Alexandros Chatzigeorgiou, Professor, University of Macedonia

____
Anastasios Oikonomidis, Professor, University of Macedonia

____
Author: Christos L. Stergiou

Thessaloniki, 2021

Acknowledgements

I would first like to thank my supervisor Prof. Dr. Konstantinos E. Psannis, for accepting me to be a member of his team Mobility2Net in the University of Macedonia. Furthermore, I am greatful for providing me with his constant support, his always detailed and immediate feedback, and his endless patience. The completion of this dissertation would not have been possible without his understanding, and guidance. Apart from sharing invaluable academic knowledge, he has shown me that being an independent scientist and at the same time an active part of a research group. During the four years of my doctoral research we had perfect cooperation and his useful tips were crucial in overcoming the various obstacles I faced.

My sincere thanks to my friends and collegues Andreas P. Plageras, Vasilis A. Memos and Dr. Katerina Tsarava, because my research would not have been possible completed without their support. Our collaboration was impeccable and often led to very good research results, which were very constructive for the continuation and completion of my research. It has been a great honor working with them. I will be happy to continue working with them in the future. Moreover, I wish to thank my friend Dr. Konstantinos Arampapaslis for his “unofficial” supervisor and the text editing, as he helped me improve my English language skills.

No words can describe my gratitude to my family. They gave me unconditional love, sacrificed a lot and supported every single idea I have had in all of these years. They helped and supported me throughout my studies and research and it was always there for me.

Last, but surely not least, I would like to deeply thank my wife Alexandra for her invaluable love and support.

Abstract

The rapid development of modern technologies such as Cloud Computing (CC), Big Data (BD), Internet of Things (IoT), Wireless Communication Systems (WCS), and Artificial Intelligence (AI) significantly affect many activities of modern everyday life. The great advancements in communication technologies and many other technology-based sectors are causing increasing security and privacy issues. Additionally, the huge hardware infrastructures that emerging technologies rely on demand large amounts of electricity highlighting the importance of energy-efficient and green infrastructures. This dissertation aims to survey the operation of CC, BD, IoT, and WCS in relation to this important issue. More specifically, I will try to explore the challenges that the integration of these technologies creates, mainly focusing on issues of security, privacy, data management, and energy-efficient use.

During my PhD, I have reported numerous findings that could offer new opportunities of having a more secure and energy-efficient environment, based on CC technology. All the research papers and manuscripts included in the present dissertation are listed in a logical order, starting from theoretical research and moving on to practical experimental studies. All the studies conducted during my PhD aimed publication at high-ranking journals and were presented in significant conferences of the broader Technology Communication field.

The studies present possible integrations of CC with technologies such as IoT and BD. The main scope was to find gaps in the secure use of BD, which most often, are produced by IoT, in cloud environments. The first chapter of this dissertation is a general introduction to all the examined technologies. The published papers and in-press manuscripts presented in chapters two, three, four, and five are examining CC, BD, IoT, and their security and privacy issues aiming to propose novel algorithms which are based on the existing encryption algorithms, offering better integration models.

The papers presented in chapters six, seven, eight, and nine examine and present novel scenarios and frameworks that use IoT-based BD, through WCN, which are based and dependant on CC. Many of the proposed scenarios and frameworks are settled and simulated on very well-known and important simulators, such as CloudSim and Cooja Contiki. The focus of these papers was to provide a better managing building system installed in a Smart Building.

The papers presented in chapters ten, eleven, and twelve are based on the idea of Smart Building and the communication system they integrate. More specifically, they offer new opportunities for managing, transferring, and processing data, in many cases IoT-Big Data, via a wireless network, based on Cloud infrastructures. Main objective of this research effort is to propose more secure environments for managing, transferring, and processing data. These simulations revealed the need for energy-

efficient infrastructures and, in some cases, the need for AI and Machine Learning (ML) methods involvement.

Finally, chapters thirteen, fourteen, and fifteen present novel scenarios of “green” infrastructures that are based on ML. As in all the aforementioned papers and manuscripts, the main scope is to provide an environment for better management, transfer, and processing of IoT-based BD. This environment mainly operates on WCN and is based on CC. The implications of my PhD research are presented more precisely in subsection 1.4.

Keywords

Cloud Computing, Big Data, Analytics, Security, Privacy, Efficiency, Internet of Things, Algorithms, Data Management, Simulation, Energy Efficiency, CloudSim, Machine Learning.

Contents

List of Figures

List of Tables

Abbreviations

ABAC	Attribute-Based Access Control
ABE	Attribute Based Encryption
AES	Advanced Encryption Standard
AF	Amplify-and-Forward
AI	Artificial Intelligence
AP	Access Point
ARC	Adaptive Replacement Cache
A-SAC	Attribute-based Semantic Access Control
ASBD	Academic - Scholarly Big Data
AVC	Advanced Video Coding
BD	Big Data
BDA	Big Data Analytics
BI	Business Intelligence
BI	Business Intelligence
BL	Base Layer
CaSF	Cloud-assisted Smart Factory
CBDM	Controllable Blockchain Data Management
CC	Cloud Computing
CCIoT-CMfg	CC- and IoT-based Cloud Manufacturing
CCo or ClCo	Cloud Coordinator
CDA	Cache Decision Algorithm
CDS	Cache Decision System
CDS	Cache Decision System
CFL	Compact Florescent Light
CMA	Comparative Market Analysis
CMfg	Cloud Manufacturing
COIB	Cognitive Oriented IoT Big-data
CoTs	Cloud of Things
CPU	Central Processing Unit
CSe	Cloud Server
CSPs	Cloud Service Providers
D2ES	Dynamic Data Encryption Strategy
D2R	Dynamic Demand Response
DAPFs	Distributed Application Processing Frameworks
DBMS	DataBase Management Systems
DCIE	Data Center’s Infrastructure Efficiency
DES	Data Encryption Standard
DF	Decode-and-Forward
DSP	Digital Signal Processor
DSS	Decision Support System
DVFS	Dynamic Voltage and Frequency Scaling
EC	Edge Computing
EEIBDM	Energy-Efficient industrial IoT-based Big Data Management Framework
EIPE	European ICT Poles of Excellence
EL	Enhancement Layer
ESAS	Extremely Scale Analytics System

ESe	Edge Server
EU	European Union
FedAvg	Federated Averaging
FedSGD	Federated Stochastic Gradient Descent
FIFO	First In First Out
FSVRG	Federated Stochastic Variance Reduced Gradient
HD	High Definition
HDR	High Dynamic Range
HEVC	High-Efficiency Video Coding
HNs	Hybrid Networks
HPN	High-Performance Network
HSTSM	Hierarchical Spatial-Temporal State Machine
HTTP	Hypertext Transfer Protocol
IaaS	Infrastructure as a Service
IBSC	Identity-Based SignCryption
ICN	Information-Centric Networking
ICT	Information and Communication Technology
IDC	International Data Corporation
IDM	Internet Download Manager
IDs	identifiers
IFC	Information Flow Control
IIoT	Industrial Internet of Things
InFeMo	Integrated Federated Model
IoT	Internet of Things
IT	Information Technology
ITEP	IT Equipment Power
ITS	Intelligent Transport Systems
JCT-VC	Joint Collaborative Team on Video Coding
KG	Key Generation System
LFRU	Least Frequent Recently Used
LFU	Least Frequently Used
LIFO	Last In First Out
LRU	Least Recently Used
M2M	Machine-to-Machine
MCC	Mobile Cloud Computing
MCNs	Mobile Cellular Networks
MOPA	Multi-Objective Privacy-Aware
MPEG	Moving Picture Experts Group
MRU	Most Recently Used
NDN	Named Data Networking
NIST	National Institute of Standards and Technology
OS	Operating System
PaaS	Platform as a Service
PCE	Pervasive Computing Environment
PDA	Proactive De-terminative Access
PSNR	Peak Signal-to-Noise Ratio
PUE	Power Usage Effectiveness
QoS	Quality of Service

RAS	Reputation Authentication System
RBT	Resource-Based Theory
RC5	Rivest Cipher 5
RFID	Radio-frequency identification
RL	Reinforcement Learning
RSA	Rivest–Shamir–Adleman
RTT	Round Trip Tim
RTUs	Remote Telemetry Units
SaaS	Software as a Service
SB	Smart Buildings
SGD	Stochastic Gradient Descent
SGX	Software Guard Extensions
SINR	Signal to Interference and Noise Ratio
SLAs	Service Level Agreements
SNg	Social Networking
SoIP	Storage over Internet Protocol
TCGs	Temporal Conceptual Graphs
TCP/IP	Transmission Control Protocol/Internet Protocol
TFP	Total Facility Power
TPM	Trusted Platform Module
TPM	Trusted Platform Module
UHD	Ultra High Definition
UML	Unified Modeling Language
VM	Virtual Machine
VMM	Virtual Machine Manager
VS	Video Surveillance
WSAN	Wireless Sensor and Actuator Network
WSN	Wireless Sensor Network

Chapter 1

Introduction

1.1 Theoretical Framework

With this dissertation, I will try to study and integrate technologies to provide a more efficient and secure cloud environment with the aim of managing, processing, and transferring the data, and more specifically BD.

BD can be referred to as a big thing in the field of modern technologies. As soon as we can decode the best use of BD we will have the opportunity to change the world completely by using all the extracted information of the data. To better understand the phenomenon of BD, we would have to find out the usage of their five major characteristics, which are widely known as five Vs of BD [1] [2]: 1) Volume, 2) Velocity, 3) Variety, 4) Veracity, 5) Value. However, in the early stages, of course, the researchers had focused and dealt with only the 3 basic characteristics of BD, which are Volume, Velocity, and Variety, but then it became clear that the characteristics of BD are 5 and not 3. Specifically, Volume of BD refers to the vast amounts of data that are generated every second, Velocity of BD refers to the speed at which the new data sets are generated and also the speed at which the data sets move around, Variety of BD refers to the various types of data that can be used, Veracity of BD refers to the messiness or trustworthiness of the data, and Value of BD refers to the worth of the data which have been extracted.

BD management could be used to customize the consistency level. More generally, everyone can perform more relaxed consistency-based replication systems on top of particular database storage systems that count on stricter transactional semantics. The customized replication and consistency enforcements could be considered a useful aspect of the applications, in which several updates might require higher integrity and some might require the higher scalability of relaxed consistency [3] [4].

Additionally, management will probably be the most difficult problem to address with BD. This is not a new problem in this field. It occurred years ago, where some scientists realized that data was distributed geographically and owned and managed by multiple entities [5] [6]. Analyzing the Data as BD and taking into account their features we can reach some conclusions. Particularly, the richness of digital data representation forbids an unveiled methodology for data collection. Data specification often focuses more on the missing data than trying to attest to every item. As regards the data volume, it is purposeless to attest every data item such as new approaches to data specification and ratification [5] [7].

Moreover, CC additionally could be used as a base technology due to its type of services for other relative to the communication field technologies [4] [8]. BD is a relative technology of the field of communications that could rely on CC. It is known from the literature that BD refers to the description of the stunning rise of data volume either of structured or unstructured form. In addition to this, the term BD describes a specific amount of data set [9] [10]. Therefore, the major problem that arises not relies on the gain of large amounts of data, but whether these data have any value or not. Hopefully, by envisaging that the companies of IT field would be able to extract information from any source, also utilize the pertinent data and analyze the data aiming to get immediate answers, we will achieve reducing cost and time, producing new products and optimizing offerings, and more intelligent decisions making [7] [8].

Also, CC could be referred to as an extremely successful example-oriented IT service. Also, CC has brought a new revolution in the way in which the computing infrastructure is used, and also, it could be extended to Database as either Service or Storage. Moreover, CC could be an omnipresent example due to its characteristics by setting up innovative applications. These applications were not currently economically feasible for traditional businesses. Thus, through scalable Data Base Management Systems (DBMS), which is CC's infrastructure critical part, could be achieved an update on intensive application workloads, such as decision support systems [11].

Furthermore, due to its unique use of cloud's environment, the providers and the customers of CC are keen to share the responsibility for security and privacy in CC environments; with the limitation however of that the sharing levels will differ for different delivery models, which in turn affect the cloud extensibility.

The following delivery data models are offered in the CC environment [8] [12] [13]: 1) SaaS, 2) PaaS, 3) IaaS. These models provide relation to software, platform, and the infrastructure as cloud services. Specifically, SaaS is the delivery model which could offer typically enabled services by providing a large number of integrated features, which could lead to less extensibility for the customers, PaaS is a delivery model which aims to enable developers to build their applications on top of the provided platforms, and IaaS is the delivery model which is the most extensible of the tree. In this delivery model, the Cloud providers must provide some basic, low-level data protection capabilities [12].

1.2 Open Issues in the Field

The need for "*cloud*" support has become inefficient due to intensive computations, mass storage, and security issues. Some examples include limited storage capacity, communication capabilities, energy, and processing. Inefficiencies like these have motivated us to find a model for the combination of CC and other technologies such as the IoT and BD. As a "base" technology, CC consolidates various technologies and applications to get the maximum capacity and performance of the existing infrastructure [14] [15] [16].

Regarding Security and Privacy issues and challenges in the field of CC, and to succeed a secure communication over the network, the encryption algorithm plays an important role. It is a valuable and fundamental tool for the protection of the data. The encryption algorithm converts the data into scrambled form by using "*a key*" and only the user has the key to decrypt the data. Regarding the researches that have been made in the field, an important encryption technique is Symmetric key Encryption. In Symmetric-key encryption, only one key is used to encrypt and decrypt the data. In this encryption technique, the most used algorithm is the AES algorithm [17] [18] [19].

AES (Advanced Encryption Standard) is the highly developed encryption standard recommended by NIST aiming to replace the older DES algorithm. Brute force attack is the only effective attack known against it, in which the attacker tries to test all the character combinations to unlock the encryption. The AES algorithm blocks ciphers. Also, AES has been carefully tested for many security applications [17] [20] [21].

The problem with security and privacy in everyday life could be solved or could be minimized by the use of BD analysis tools and services. BD is a new popular term, used to describe the surprisingly rapid increase in the volume of data in the structured and unstructured form [7]. Accuracy in big data may lead to more confident decisions making, and better decisions can result in greater operational efficiency, cost reduction, and reduced risk [10] [22]. BD usually uses CC as a base technology to operate.

In addition to this, CC could be used as a base technology for another relative to communications technology, IoT. The basic idea of the IoT is the diffuse presence of a variety of things or objects used by people such as radio-frequency identification tags, sensors, actuators, and mobile phones. Through unique addressing schemes, these things interact with each other and cooperate with other things near them to reach the common goals [23] [24]. The IoT can be defined as "*the network of physical objects, devices, vehicles, buildings and other items which are embedded with electronics, software, sensors, and network connectivity, permitting these objects to gather and interchange data*" regarding the bibliography [8] [25] [26]. Some examples include the restrictions of storage, communication capabilities, energy, and processing offered to

IoT devices. Those inefficiencies motivate us to combine the functionality of CC and IoT technologies [8] [27] [28].

IoT security is the area of strive concerned with safeguarding connected devices and networks in the IoT. The IoT involves the raising dominance of objects and entities, provided with unique identifiers and the ability to automatically transmit data over a network. Much of the increase in IoT communication comes from computing devices and the embedded sensor systems used in sectors such as industrial machine-to-machine (M2M) communication, smart energy grids, home and building automation, vehicle to vehicle communication, and wearable computing devices [29] [30] [31] [32].

Furthermore, the new technology called CC could be defined as "*a distributed information technology (IT) architecture in which client data is processed at the periphery of the network, as close to the originating source as possible*" [33] [34] [35]. The move toward Cloud Computing is driven by mobile computing, the decreasing cost of computer components, and the absolute number of networked devices in the IoT. More specifically, CC refers to data processing power in a fog network instead of holding that processing power in a cloud or a central data warehouse [33] [36] [37] [38].

CC storage solutions offer users and enterprises various capabilities to store and process their data in third-party data centers [1] [34] [39] [40] [41]. To offer safer communication over the network, the encryption algorithm plays a vital role. It is a valuable and fundamental tool for the protection of the data. The encryption algorithm converts the data into scrambled form by using "*a key*" and only the user has this key to decrypt the data. Regarding the researches which have been carried out, an important encryption technique is Symmetric key Encryption. In Symmetric-key encryption, only one key is used to encrypt and decrypt the data. In this encryption technique the most used algorithm is the AES [1] [42] [43] [44] [45].

Internet of Things characteristics	*Storage over Internet*	*Service over Internet*	*Applications over Internet*	*Energy efficiency*	*Computationally capable*
Smart solution in the bucket of transport	X	X	X		X
Smart power grids incorporating more renewable	X	X		X	X
Remote monitoring of patients		X	X		X
Sensors in homes and airports	X	X	X	X	X
Engine monitoring sensors that detect & predict maintenance issues		X	X	X	X

Table 1.1: Contributions of Cloud Computing in the Internet of Things [80].

Table 1.1 lists the basic characteristics of CC technology as regards the convenience this technology offers. Also, it enumerates the basic characteristics of IoT technology. The main purpose of Table 1 is to demonstrate which of the specific characteristics of CC technology, related more and improve the characteristics of IoT technology. As we could realize by observing Table 1, the characteristics of IoT that are affected more by the characteristic of CC is "*Sensors in homes and airports*". Regarding the CC, the characteristics that affected more are "*Service over Internet*" and "*Computationally capable*". As a general conclusion, we can observe that those two technologies contribute more to each other in many of their characteristics.

Moreover, as regards the data used in a wireless network there are security and privacy issues that need to be addressed. The problem with security and privacy in everyday life could be solved or could be minimized by the use of BD analysis tools and services. BD is a novel popular term, used to derive the surprisingly rapid increase in the volume of data in any form (structured and unstructured) [46] [23] [47]. BD usually uses CC as a base technology to operate. Similar to this, another technology that could be used as a base technology is Edge Computing (EC).

In addition to this, CC and EC could be used as a base technology for multiple affiliated technologies of the communications field, such as IoT that mentioned before. The basic idea of the IoT relays to the pervasive presence of various things or objects used by people such as radio-frequency identification tags, sensors, actuators, and mobile phones. Through exclusive addressing schemes, these things communicate with each other and cooperate with other things near them aiming to reach the common goals [8] [48] [49]. Thus, the exhaustive computations and the mass storage, which are supported by clouds, are sometimes inefficient. Several examples contain the limitations of storage, communication capabilities, energy, and processing.

Inefficiencies such as these motivate us to combine the technology of CC, EC, and IoT [50] [51] [52].

Therefore, another real issue in the field of managing and transferring BD in the Cloud is not that large amounts of data have to be acquired, but whether it has any value or not. Hopefully, by envisaging that the organizations would be able to acquire information from any source, harness the relevant data and analyze it to get quick answers, I will try to achieve the following: 1) reduce costs, 2) reduce time, 3) produce new commodities and to optimize their offerings, 4) make more intelligent decisions.

Furthermore, in the last years, there is a standard of High-Efficiency Video Coding, known as HEVC, which is the latest compression standard. This type of video compression standard could be used and transmitted as data sets, which could be defined as Smart Big Data. HEVC was officially approved in January 2013 and became the successor of the H.264/MPEG-4 or AVC standard. The new technology is called HEVC also known as H.265 and MPEG-H Part2. Compared to AVC video coding the new technique of video coding, HEVC provides about two times the data compression ratio at the proper level of video quality or essentially meliorated video quality at the same bit-rate.

Moreover, the HEVC video coding patronizes resolutions up to 8192x4320, including the 8K UHD. The data used by the HEVC could also be characterized as BD due to their large volume [24]. The basic goal of the HEVC standard is the circumstance that there is a presentation of considerably better compression performance in contrast with the current existing standards. This could be in the range of 50% bit-rate reduction in nearly the same video quality, compared to H.264/MPEG-AVC standard [53]. Thus, the HEVC was a design created to offer high-quality streaming media, even on low-bandwidth networks. As a result, it consumes only half bandwidth compared to AVC [54]. Cooperation between the ISO/IEC MPEG and ITU-T VCEG has the result that the JCT-VC organization developed the HEVC. The ISO/IEC group refers to it as MPEG-H Part 2 and the ITU-T as H.265. In January of 2013, the first version of HEVC was created and published in June 2013. In 2014, the second version was completed and approved and then published at the beginning of 2015. In addition to this, 3D-HEVC expansions for the 3D video were completed at the beginning of 2015.

Finally, CC is a technology that counts on huge data centers, and as a result, counts on enormous amounts of electricity. Conventional Cloud applications are offered from social networking, content delivery, web hosting, data processing, and others. Such applications have multiple and different requirements. Also, the volumes of the system and power performance are two open challenges of these applications on Cloud Service Providers. Cloud Service Providers (CSPs) use their infrastructures to provide all the available Cloud models (IaaS, PaaS, and SaaS) [49]. Energy amounts of such CSPs data centers are rising every year due to users' demands. Cloud federation is a novel architecture that might offer possibilities and opportunities to address the open challenges.

Moreover, to achieve better Cloud usage a lot of scenarios are proposed by the researchers. Cloud simulators are the most efficient method to facilitate and operate a Cloud infrastructure without the additional costs of the hardware required. There are many Cloud Computing simulators, which are simulating many aspects of Cloud environments. The most widely known are CloudSim, CloudAnalyst, GreenCloud, NetworkCloud, EMUSIM, and MDCSim [7] [55]. In this dissertation, I used the CloudSim simulator to set up and evaluate our proposed scenario which relying on Green Cloud infrastructure.

1.3 Problem Definition

CC offers abilities and functions such as computing, storage, services, and applications over the Internet. In general, to render smartphones energy-efficient and computationally capable, major changes to the hardware and software levels are required. This causes the cooperation of developers and manufacturers [56].

1.3.1 Cloud Computing Features

Like all technologies, so CC technology has several characteristics which determine its operation. These characteristics are represented and outlined below.

***CC(a):** Storage over Internet*

Storage over Internet can be defined as "*a technology framework that uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks to link servers and storage devices and to facilitate storage solution deployment*" [57] [58].

***CC(b):** Service over Internet*

The Service over Internet has as a major objective is to "*help customers all over the world to transform aspirations into achievements by harnessing the Internet's efficiency, speed and ubiquity*" [57] [58].

***CC(c):** Applications over Internet*

Cloud Applications, or as scientific known as Applications over Internet, are the programs that have been written to do the job of a current manual task, or virtually anything, and which perform their job on the server through an internet connection [57] [58].

***CC(d):** Energy Efficiency*

Energy Efficiency could be defined as "*a way of managing and restraining the growth in energy consumption*" [57] [58]. By delivering more services for the same energy input or the same services for less energy input maybe something more energy-efficient [57] [58].

***CC(e):** Computationally Capable*

The services of computational clouds are leveraging the computationally concentrated and ubiquitous mobile applications which have been enabled by the technology of MCC. Thus, a system can be considered computationally capable when it meets the requirements to offer us the results we want, by making the right calculations [57] [58].

1.3.2 Security on Cloud Computing

CC security is an evolving sub-domain of computer security, network security, and information security. It alludes to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of CC.

CC technology offers through its storage solutions to users and industries various capabilities to store and process their data in third-party data centers [39]. Thus, by aiming to offer secure communication through the network, the encryption algorithm plays a vital role. As regards the researches that have been made, an important encryption technique is Symmetric Key Encryption. In Symmetric-key encryption, only one key is used to encrypt and decrypt the data. In this encryption technique, the most used algorithm is the AES [42] [43] [59].

1.3.3 Cloud Computing trade-offs

Cloud Computing has some disadvantages-limitations which should be eliminated over the years to achieve a better and more ideal use. Some businesses and especially the smaller ones need to be aware of these limitations before going in for this technology.

***CC(l-a):** Security*

One major issue of the MCC is the security issue. Before someone adopts this technology, they should know that all the company’s sensitive information would be surrender to a third-party cloud service provider. This could potentially put the company at great risk. Hence, someone must be sure that they would choose the most reliable service provider, who will keep the information completely safe [60] [61] [62].

***CC(l-b):** Connectivity*

Internet connection is critical to CC. Thus, the user should be certain that there is a good result before opting for these services. Since someone owes a mobile device that is connected to the internet has become the norm in the wireless world of today, CC has a very large potential user base [60] [63].

CC(l-c): *Performance*

Another major concern of the CC pertains to its performance. Some users feel performance is not as good as in native applications. Thus, checking with one service provider and understanding their track record is advisable [60] [64].

CC(l-d): *Latency (Delay)*

In CC, latency (sometimes referred to as turnaround time) is defined as the time involved in offloading the computation and getting back the results from the nearby infrastructure or cloud [60] [65].

CC(l-e): *Privacy*

Data privacy is important and is one of the main bottlenecks that restrict consumers from adopting CC. Therefore, to gain consumers' trust in the Cloud, the application models must support application development with privacy protection, and implicit authentication mechanisms [60] [62] [66].

1.3.4 Challenges outcomes by combining Cloud Computing with other technologies

When critical applications, such as the IoT applications, move towards the CC technology, concerns arise due to the lack of trust in the service provider or the knowledge about service level agreements (SLAs) and knowledge about the physical location of data. Consequently, new challenges require specific attention as mentioned in surveys [67] [68] [69]. Multi-tenancy could also compromise security and lead to sensitive information leakage. Moreover, public-key cryptography cannot be applied at all layers due to the computing power constraints imposed by the things. These are examples of topics that are currently under investigation to tackle the big challenge of security and privacy in integrating CC with other technologies [67].

Subsequently, some challenges about the security issue in the integration of CC with other technologies are listed below [67].

a) Heterogeneity: A big challenge in CC integration with other technologies is related to the wide heterogeneity of devices, operating systems, platforms, and services available and possibly used for new or improved applications [70].
b) Performance: Often CC integration with other technologies applications introduce specific performance and QoS requirements at several levels (i.e. for communication, computation, and storage aspects), and in some particular scenarios meeting requirements may not be easily achievable [71]
c) Reliability: When CC integration with other technologies is adopted for mission-critical applications, reliability concerns typically arise. When applications are deployed in resource-constrained environments several challenges related to device failure or not always reachable devices exist [72].
d) Big Data: With an estimated number of 50 billion devices that will be networked by 2020, specific attention must be paid to transportation, storage, access, and processing of the huge amount of data they will produce [73].

e) Monitoring: As largely documented in the literature, monitoring is an essential activity in Cloud environments for capacity planning, for managing resources, SLAs, performance, and security, and troubleshooting [74].

1.3.5 Cloud Computing as a base technology for Big Data

The recent years, CC services compose one of the major areas in the world of competition among the giant companies in the field of IT and software [75] [76].

More specifically, CC is consisted of technology of internet services providing remote use of hardware and software. As a result, the users of CC could have access to information and data from any place at any time. In the concept of CC, there is another technology called MCC that refers in general concept to two prospects [77]: a) infrastructure-based, and b) ad-hoc mobile cloud. In the prospect of an infrastructure-based mobile cloud, the infrastructure of the hardware is still static and delivers services to the mobile users [78] [79]. Particularly, MCC is defined as "*the integration of CC and Mobile technology to make any type of mobile devices resourceful in terms such as computational power, memory, storage and energy*" [81]. Regarding the usage of Cloud services in Mobile devices, many types of services could be processed through it. Thus, high-quality media could be transmitted through the Cloud environment progressed in applications that were installed and operated in Cloud.

Considering this CC could be settled as a base technology to operate other technologies such as BD and consequently to be accomplished integration of Cloud and Big Data [58] [83]. In addition to this, CC also used to be a base technology for other technologies due to its type of services [4] [80].

As already mentioned, one of those is BD. It is used to describe the surprisingly rapid increase in the volume of data in structured and unstructured forms. It is a broad term for data sets so large or complex that traditional data processing applications are inadequate. Furthermore, BD often refers to the use of predictive analytics or certain advanced methods to extract value from data. Rarely, it also refers to a particular size of the data set [9] [10]. Precision in BD could result in more confident decision making, and better decisions may lead to increased operational efficiency, reduced costs, and minimized risk [9]. From this scope, I realize that BD is now equally important both for business and the internet. This happens because more information leads to more accurate analyses [4]. The real problem is not that you have acquired large quantities of data, but whether it has any value or not. Hopefully, by envisaging that the organizations would be able to obtain information from any source, harness the relevant data and analyze it to get quick answers, we will achieve the following: 1) to reduce costs, 2) to reduce time, 3) to produce new products and to optimize their offerings, 4) to make more intelligent decisions [1] [7].

Moreover, another novel technique that offers new opportunities to manage and operate the data in cooperative environments makes its appearance. This novel technique is called federated learning. Thus, according to the literature on federated learning, the main objective is to train a model from data $\{X^t,\ldots,X^K\}$ produced by K distributed clients. Every device is represented as a client. Each client, $t \in [K]$, produces data in a Non-IID manner, which means the data distribution on the client t, $X^t \sim P^t$, is not a uniform sample of the whole distribution [81]. Based on the literature, the federated learning technique bases on distributed machine learning to which a global model is learned by aggregating models that have been trained locally on data-generating clients [82]. Additionally, the algorithms of federated learning scenarios reckon with the fact that communication with edge devices occurs over unreliable networks with very limited upload speeds [82]. As a result, federated learning can significantly decrease the privacy and security risks by limiting the attack surface only to the device, and not to both the device and the Cloud [83]. Particularly, Federated learning tries to give solutions to problems such as 1) The distinct advantage of training on proxy data that is generally vacant in the data center could be provided by training on real-world data delivered by mobile devices. 2) Trained data is better not to attain it to the data center wholly for the intention of model training, due it is privacy-sensitive or large. 3) User interaction could infer natural labels on the data for supervised tasks [83].

The optimization problem of federated learning could be defined with an algorithm that optimizes the finite-sum objective:

$$\min_{w \in R^d} f(w) \quad \text{where} \quad f(w) = \frac{1}{n}\sum_{i=1}^{n} f_i(w) \quad \textbf{(1)}$$

In equation (1) the w is a vector that contains d model parameters. In machine learning scenarios, we treat the function $f_i(w)$ as a loss function $f_i(w) = \ell(x_i, y_i; w)$, where an input-output pair (x_i, y_i) is one of the n given labeled examples, in most times referred to as a *training example*. One problem could be interpreted as finding the w which minimizes the average loss over all n training examples such as those in [84]. Moreover, another scenario is assuming that there are K clients over which the data is partitioned, with P_k the set of indexes of data points on the client k, $n_k = |P_k|$. This could lead us to produce a new equation from (1) [85]. Furthermore, regarding the BD context, which is the main technology apart of the CC I focus on my research, I can state that the number of training examples is too large to be stored on one computer, and thus I need to distribute the computation to many computers. Concluding these we could have:

$$f(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w) \quad \text{where} \quad F_k(w) = \frac{1}{n_k}\sum_{i \in P_k} f_i(w) \quad \textbf{(2)}$$

From equation (1) and (2) could be extracted the federated learning parameters, such as 1) the number of federated learning rounds, represented as T, 2) the total number of nodes used in the process, represented as K, 3) the fraction of nodes that used at each iteration for each node, represented as C, 4) the local batch size which used at each learning iteration, represented as B, 5) the number of iterations for local training before the pooling, represented as N, and 6) the local learning rate, represented as η. These parameters need to be optimized by the constraints of the machine learning applications. Also, in addition to the above parameters, the main strategies of federated learning could be described with some notations, such as the K, which is the total number of clients, the k, which illustrates the index of clients, n_k, which refers to several data samples usable during training for client k, $w_{k,t}$ which demonstrates the model's weight vector on the client k, at the federated round t, $l(w,b)$, which illustrates the loss function for weights w and batch b, and E, which represents the number of local epochs.

Thereafter, there is also the Federated Stochastic Gradient Descent, or widely known as FedSGD of SGD, which is a deep learning training primarily count on variants of stochastic gradient descent, at which place gradients are computed on a random subset of the total dataset and afterward used to make one step of the gradient descent. FedSGD which was initially proposed by Shokri & Shmatikov [85] is the direct transposition of the FedSGD algorithm to the federated setting. However, the gradients are averaged by the server proportionally to the number of training samples on each node and used to make a gradient descent step.

1.4 Thesis's Contributions

This dissertation based on the research finding and the open issues in the field of Big Data management, processing, and transfer through a Cloud environment tries to offer some innovative solutions and proposals keen on the upcoming contributions.

- ✓ Examines CC, BD, and other relative technologies such as IoT.
- ✓ Considers *CloudSim*'s simulator stable architecture.
- ✓ Presents a sort survey of IoT and CC with a focus on the security issues of both technologies.
- ✓ Integrates IoT and CC intending to examine the common features, and to discover the benefits of their integration
- ✓ Proposes an efficient algorithm for advanced scalable Media-based Smart Big Data (3D, Ultra HD) on Intelligent CC systems. Tries to conclude that this proposed method could be used and integrated into HEVC, as a Smart BD, without violating the standard.
- ✓ Proposes an innovative system of secure caching scenario which operates in a wireless-mobile 6G network for managing BD on Smart Buildings (SB). The proposed scenario combines the functions of the IoT with CC, EC, and BD.
- ✓ Creates a novel and secure Cache Decision System (CDS) in a wireless

network that operates over a SB, which offers the users a safer and efficient environment for browsing the internet, sharing, and managing large-scale data in the fog. The CDS model consists of two types of servers, one Cloud Server (CSe) and one Edge Server (ESe).

- ✓ Investigates and proposes a system framework for better use of BD management, based on a Cloud federated network.
- ✓ Proposes an algorithm for achieving an energy-efficient resource allocation technique for BD management.
- ✓ Provides a more energy-efficient system architecture and environment for the academic users with the aim of data management.
- ✓ Decreases number of rounds of communication needed to train a scenario model by using a federated Cloud system. Thus the users have to wait for less. This tries to fill a scientific gap in the field of federated cloud systems management.
- ✓ Proposes an innovative architecture model, *InFeMo – Integrated Federated Model*, that incorporates all Cloud models with a federated learning scenario, as well as other technologies that may have integrated use with each other, in a novel integrated scenario.

1.5 Outline

The dissertation is structured in sixteen chapters. Chapters two to fifteen list and illustrate the major and important researches that have been made in the context of my doctoral research. Each chapter from two to fifteen is organized as follows: Introduction, Basic Research Item, Research Findings, and Chapter Summary (as Conclusion). Specifically, chapter two presents the work titled "*Algorithms for Big Data in Advanced Communication Systems and Cloud Computing*", chapter three presents the work titled "*Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey*", chapter four presents the work titled "*Secure integration of IoT and Cloud Computing*", chapter five presents the work titled "*Security & Privacy of IoE-based Big Data in Cloud Computing*", chapter six presents the work titled "*Recent advances delivered in Mobile Cloud Computing's Security and Management challenges*", chapter seven presents the work titled "*Secure Machine Learning scenario from Big Data in Cloud Computing via Internet of Things network*", chapter eight presents the work titled "*Efficient and Secure Big Data delivery in Cloud Computing*", chapter nine presents the work titled "*Algorithms for efficient digital media transmission over IoT and cloud networking*", chapter ten presents the work titled "*Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT*", chapter eleven presents the work titled "*IoT-based Big Data secure management in the Fog over a 6G Wireless Network*", chapter twelve presents the work titled "*Advanced Media-based Smart Big Data on Intelligent Cloud Systems*", chapter thirteen presents the work titled "*Green Cloud Communication System for Big Data Management*", chapter fourteen presents the work titled "*InFeMo: Flexible Big Data management through a federated Cloud* *system*”, and chapter fifteen presents the work titled “*EEIBDM: A Reinforcement & Federated Learning scenario for Efficient Industrial IoT-based Big Data Management in Cloud*”. Chapter sixteen summarizes the main contributions of this dissertation and outlines the basic steps of future plans.

1.6 Publications

All the important research findings of this dissertation are published in peer-reviewed journals and conference papers. All papers underwent extensive reviewing processes before the actual publication to the scientific community. The complete list of publications is presented below.

Peer-reviewed journals:

- **C. Stergiou**, K. E. Psannis, “Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey”, Wiley, International Journal of Network Management, pp. 1-12, May 2016. [DOI:10.1002/nem.1930]
- **C. Stergiou**, K. E. Psannis, “Efficient and Secure Big Data delivery in Cloud Computing”, Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017. [DOI:10.1007/s11042-017-4590-4]
- **C. Stergiou**, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI:10.1016/j.future.2016.11.031]
- **C. Stergiou**, K. E. Psannis, A. P. Plageras, Y. Ishibashi, B.-G. Kim, “Algorithms for efficient digital media transmission over IoT and cloud networking”, Journal of Multimedia Information System, vol. 5, issue: 1, pp. 27-34, March 2018. [DOI: 10.9717/JMIS.2018.5.1.27]
- **C. Stergiou**, K. E. Psannis, B. Gupta, Y. Ishibashi, “Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT”, Elsevier, Sustainable Computing, Informatics and Systems, vol. 19, pp. 174-184, September 2018.
- K. E. Psannis, **C. Stergiou**, B. B. Gupta, “Advanced Media-based Smart Big Data on Intelligent Cloud Systems”, IEEE Transaction on Sustainable Computing, vol. 4, Issue: 1, pp. 77-87, January-March 2019.
- **C. L. Stergiou**, K. E. Psannis, B. B. Gupta, “InFeMo: Flexible Big Data management through a federated Cloud system”, ACM Transactions on Internet Technology, In Press, 2020. [DOI: 10.1145/3426972]
- **C. L. Stergiou**, K. E. Psannis, B. B. Gupta, “IoT-based Big Data secure management in the Fog over a 6G Wireless Network”, IEEE Internet of Things Journal, In Press, 2020. [DOI: 10.1109/JIOT.2020.3033131]

Peer-reviewed journal papers under review:

- **C. L. Stergiou**, K. E. Psannis, "Security & Privacy of IoE based Big Data in Cloud Computing", IEEE Internet of Things Journal (*to be submitted*)
- **C. L. Stergiou**, K. E. Psannis, "EEIBDM: A Reinforcement & Federated Learning scenario for Efficient Industrial IoT-based Big Data Management in Cloud", IEEE Transactions on Sustainable Computing (*submitted 11/02/2021*)

Peer-reviewed book chapter papers:

- **C. Stergiou**, A. P. Plageras, K. E. Psannis, B. B. Gupta, "Secure Machine Learning scenario from Big Data in Cloud Computing via Internet of Things network", Springer, Handbook of Computer Networks and Cyber Security: Principles and Paradigms, Multimedia Systems and Applications, 2019.
- **C. Stergiou**, K. E. Psannis, "Recent advances delivered in Mobile Cloud Computing's Security and Management challenges", IGI Global, Modern Principles, Practices, and Algorithms for Cloud Security, 2019.

Peer-reviewed conference papers:

- **C. Stergiou**, K. E. Psannis, "Algorithms for Big Data in Advanced Communication Systems and Cloud Computing", in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece. [DOI: 10.1109/CBI.2017.28]
- **C. L. Stergiou**, K. E. Psannis, Y. Ishibashi, "Green Cloud Communication System for Big Data Management", in Proceedings of The 3rd World Symposium on Communication Engineering (WSCE 2020), 9-11 October 2020, held Online, Thessaloniki, Greece. [DOI 10.1109/WSCE51339.2020.9275579]

Peer-reviewed papers related to the main scope of the dissertation but not part of the basic research of it listed below. Full papers are represented in the Appendix of this dissertation.

Peer-reviewed journals:

- A. P. Plageras, K. E. Psannis, **C. Stergiou**, H. Wang, B. B. Gupta, "Efficient IoT-based sensor BIG Data collection-processing and analysis in Smart Buildings", Future Generation Computer Systems, vol. 82, pp. 349-357, May 2018. [DOI: 10.1016/j.future.2017.09.082]
- K. Psannis and **C. Stergiou**, "Effective and secure transfer of BIG Data to Cloud Computing with an algorithm", Journal-Periodical Edition of the University of Macedonia (T16), August 2018.

Peer-reviewed conference papers:

- **C. Stergiou**, K. E. Psannis, A. P. Plageras, G. Kokkonis, Y. Ishibashi, “Architecture for Security in IoT Environments”, in Proceedings of 26th IEEE International Symposium on Industrial Electronics, 19-21 June 2017, Edinburgh, Scotland, UK. [DOI: 10.1109/ISIE.2017.8001447]
- A. P. Plageras, **C. Stergiou**, K. E. Psannis, G. Kokkonis, Y. Ishibashi, Byung-Gyu Kim, Brij Gupta, “Efficient Large-Scale Medical Data (eHealth Big Data) Analytics in Internet of Things”, in Proceedings of 19th IEEE International Conference on Business Informatics (CBI’17), International Workshop on the Internet of Things and Smart Services (ITSS2017), 24-26 July 2017, Thessaloniki, Greece. [DOI: 10.1109/CBI.2017.3]
- A. P. Plageras, **C. Stergiou**, K. E. Psannis, Byung-Gyu Kim, Brij Gupta, Y. Ishibashi, “Solutions for Inter-connectivity and Security in a Smart Hospital Building”, in Proceedings of 15th IEEE International Conference on Industrial Informatics (INDIN 2017), 24-26 July 2017, Emden, Germany. [DOI: 10.1109/INDIN.2017.8104766]
- **C. Stergiou**, K. E. Psannis, A. P. Plageras, T. Xifilidis, B. B. Gupta, “Security and Privacy of Big Data for Social Networking Services in Cloud”, in Proceedings of IEEE conference on Computer Communications (IEEE INFOCOM 2018), 15-20 April 2018, Honolulu, HI, USA.
- **C. Stergiou**, A. P. Plageras, K. E. Psannis, T. Xifilidis, G. Kokkonis, S. Kontogiannis, K. Tsarava, A. Sapountzi, “Proposed High Level Architecture of a Smart Interconnected Interactive Classroom”, in Proceedings of IEEE conference SEEDA-CECNSM 2018, 22-24 September 2018, Kastoria, Greece.
- A. P. Plageras, **C. L. Stergiou**, K. E. Psannis, “Internet of Things for Healthcare: Challenges & Perspectives”, in Proceedings of New Technologies in Health: Medical, Legal & Ethical Issues, 21-22 November 2019, Thessaloniki, Greece.
- V. A. Memos, G. Minopoulos, **C. Stergiou**, K. E. Psannis, Y. Ishibashi, “A Revolutionary Interactive Smart Classroom (RISC) with the Use of Emerging Technologies”, in Proceedings of 2nd International Conference on Computer Communication and the Internet (ICCCI 2020), 26-28 June 2020, Nagoya Institute of Technology, Japan. [DOI: 10.1109/ICCCI49374.2020.9145987]

1.7 Chapter References

[1] C. Stergiou, K. E. Psannis, “Algorithms for Big Data in Advanced Communication Systems and Cloud Computing”, in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece. [DOI: 10.1109/CBI.2017.28]

[2] B. Marr, “Big Data: The 5 Vs Everyone Must Know”, LinkedIn article, 6 March 2014. Retrieved: 17/12/2018. Link: https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know

[3] Z. Lv, A. K. Singh, “Big Data Analysis of Internet of Things System”, ACM Transactions on Internet Technology, vol. 0, issue: ja, Accepted on March 2020. [DOI: 10.1145/3389250]

[4] C. Stergiou, K. E. Psannis, “Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey”, Wiley, International Journal of Network Management, pp. 1-12, May 2016. [DOI:10.1002/nem.1930]

[5] M. M. Rathore, A. Paul, A. Ahmad, M. Anisetti, G. Jeon, “Hadoop-Based Intelligent Care System (HICS): Analytical Approach for Big Data in IoT”, ACM Transactions on Internet Technology, vol. 18, issue: 1, No. 8, 24 pages, November 2017. [DOI: 10.1145/3108936]

[6] H. Yu, J. Yang, C. Fung, “Fine-grained Cloud Resource Provisioning for Virtual Network Function”, IEEE Transactions on Network and Service Management, vol. 17, issue: 3, pp/ 1363-1376, September 2020.

[7] C. Stergiou, K. E. Psannis, “Efficient and Secure Big Data delivery in Cloud Computing”, Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017. [DOI:10.1007/s11042-017-4590-4]

[8] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI:10.1016/j.future.2016.11.031]

[9] M. Hilbert, P. López, “The World’s Technological Capacity to Store, Communicate, and Compute Information”, AAAS, Science, vol. 332, issue: 6025, pp. 60–65, April 2011. [DOI:10.1126/science.1200970]

[10] Z. Fu, K. Ren, J. Shu, X. Sun, F. Huang, “Enabling Personalized Search over Encrypted Outsourced Data with Efficiency Improvement”, IEEE Transactions on Parallel and Distributed Systems, vol. 27, issue: 9, September 2016. [DOI: 10.1109/TPDS.2015.2506573]

[11] D. Agrawal, S. Das, A. El Abbadi, “Big Data and Cloud Computing: Current State and Future Opportunities”, pp. 530-533, in Proceedings of 14th International Conference on Extending Database Technology, EDBT 2011, 21-24 March 2011, Uppsala, Sweden.

[12] C. Pahl, P. Jamshidi, O. Zimmermann, “Architectural Principles for Cloud Software”, ACM Transactions on Internet Technology, vol. 18, issue: 2, No. 17, 23 pages, February 2018. [DOI: 10.1145/3104028]

[13] N. Ferry, F. Chauvel, H. Song, A. Rossini, M. Lushpenko, A. Solberg, “CloudMF: Model-Driven Management of Multi-Cloud Applications”, ACM Transactions on Internet Technology, vol. 18, issue: 2, No. 16, 23 pages, January 2018. [DOI: 10.1145/3125621]

[14] D. Giusto, A. Iera, G. Morabito, L. Atzori (Eds.), “The Internet of Things”, Springer, 20th Tyrrhenian Workshop on Digital Communications, 2010. ISBN: 978-1-4419-1673-0.

[15] M. Swan, “Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0”, MDPI, Journal of Sensor and Actuator Networks, vol.1, no. 3, pp. 217–253, November 2012. [DOI:10.3390/jsan1030217]

[16] J. M. Batalla, P. Krawiec, “Conception of ID layer performance at the network level for Internet of Things”, Springer, Journal Personal and Ubiquitous Computing, Vol.18, Issue 2, pp 465-480, April 2013. [DOI: 10.1007/s00779-013-0664-0]

[17] P. Bahl, Y. R. Han, L. E. Li, M. Satyanarayanan, “Advancing the State of Mobile Cloud Computing”, in Proceedings of the third ACM workshop on Mobile cloud computing and services MCS’12, pp. 21-28, 25 June 2012, Low Wood Bay, Lake District, UK. [DOI: 10.1145/2307849.2307856]

[18] O. Niggemann, J. R. Kinnebrew, H. Khorasgani, S. Volgmann, A. Bunte, G. Biswas, “Data-Driven Monitoring of Cyber-Physical Systems Leveraging on Big Data and the Internet-of-Things for Diagnosis and Control”, in Proceedings of 26th International Workshop on Principles of Diagnosis, pp. 185-192, 31 August – 3 September 2015, Paris, France.

[19] R. P. Minch, “Location Privacy in the Era of the Internet of Things and Big Data Analytics”, 48th Hawaii International Conference on System Sciences, 2015, Boise State University.

[20] R. Buyya, C. Shin Yeo, S. Venugopal, J. Broberg, I. Brandic, “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility”, Elsevier, Future Generation Computer Systems, vol. 25, Issue: 6, pp. 599-616, June 2009. [DOI: 10.1016/j.future.2008.12.001]

[21] M. Strohbach, H. Ziekow, V. Gazis, N. Akiva, “Towards a Big Data Analytics Framework for IoT and Smart City Applications”, Book Chapter, Springer,

Modeling and Processing for Next-Generation Big-Data Technologies, vol. 4, pp. 257-282, 2015. [DOI: 10.1007/978-3-319-09177-8_11]

[22] T. Li, J. Li, Z. Liu, P. Li, C. Jia. "Differentially Private Naive Bayes Learning over Multiple Data Sources", Elsevier, Information Sciences, vol. 444, pp. 89-104, May 2018.

[23] C. Stergiou, K. E. Psannis, A. P. Plageras, G. Kokkonis, Y. Ishibashi, "Architecture for Security in IoT Environments", in Proceedings of 26th IEEE International Symposium on Industrial Electronics, 19-21 June 2017, Edinburgh, Scotland, UK. [DOI: 10.1109/ISIE.2017.8001447]

[24] J. M. Batalla, P. Krawiec, A. Beben, P. Wisniewski, A. Chydzinski, "Adaptive Video Streaming: Rate and Buffer on the Track of Minimum Rebuffering", IEEE Journal on Selected Areas in Communications, vol. 34, issue: 8, pp. 2154-2167, August 2016.

[25] X. Zhang, Y.-a. Tan, C. Liang, Y. Li, J. Li, "A Covert Channel over VoLTE via Adjusting Silence Periods", IEEE Access, vol. 5, pp. 9292-9302, February 2018.

[26] Q. Lin, J. Li, Z. Huang, W. Chen, J. Shen, "A short linearly homomorphic proxy signature scheme", IEEE Access, vol. 6, pp. 12966-12972, February 2018.

[27] A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, B. B. Gupta, "Efficient IoT-based sensor BIG Data collection-processing and analysis in Smart Buildings", Elsevier, Future Generation Computer Systems, vol. 82, pp. 349-357, May 2018. [DOI: 10.1016/j.future.2017.09.082]

[28] C. Yuan, X. Li, Q. M. J. Wu, J. Li, X. Sun, "Fingerprint Liveness Detection from Different Fingerprint Materials Using Convolutional Neural Network and Principal Component Analysis", Tech Science Press, CMC: Computers, Materials & Continua, vol. 53, no. 3, pp.357-371, 2017. [DOI: 10.3970/cmc.2017.053.357]

[29] M. Rouse, "IoT security (Internet of Things security)", IoT Agenda, 01/11/2015. [Online]. Available: http://internetofthingsagenda.techtarget.com/definition/IoT-security-Internet-of-Things-security. [Accessed 27/07/2017].

[30] J. Li, X. Chen, X. Huang, S. Tang, Y. Xiang, M. M. Hassan, A. Alelaiwi, "Secure Distributed Deduplication Systems with Improved Reliability", IEEE Transactions on Computers, vol. 64, issue: 12, pp. 3569-3579, December 2015. [DOI: 10.1109/TC.2015.2401017]

[31] Q. Lin, H. Yan, Z. Huang, W. Chen, J. Shen, Y. Tang, “An ID-based linearly homomorphic signature scheme and its application in blockchain”, IEEE Access, vol. 6, pp. 20632-20640, February 2018.

[32] L. Fan, X. Lei, N. Yang, T. Q. Duong, G. K. Karagiannidis, “Secrecy Cooperative Networks With Outdated Relay Selection Over Correlated Fading Channels”, IEEE Transactions Vehicular Technology, vol. 66, no. 8, pp. 7599-7603, August 2017.

[33] M. Zakarya, L. Gillam, “Energy efficient computing, clusters, grids and clouds: A taxonomy and survey”, Elsevier, Sustainable Computing: Informatics and Systems, vol. 14, pp. 13-33, June 2017.

[34] A. Bianco, R. Mashayekhi, M. Meo, “On the energy consumption computation in Content Delivery Netwroks”, Elsevier, Sustainable Computing: Informatics and Systems, vol. 16, pp. 56-65, December 2017.

[35] J. Li, Y. K. Li, X. Chen, P. P. C. Lee, W. Lou, “A Hybrid Cloud Approach for Secure Authorized Deduplication”, IEEE Transactions on Parallel and Distributed Systems, vol. 26, issue: 5, pp. 1206-1216, May 2015.

[36] A. M. Sampaio, J. G. Barbosa, “Chapter Three – Energy-Efficient and SLA-Based Resource Management in Cloud Data Centers”, Elsevier, Advances in Computers, vol. 100, pp. 103-159, 2016. [DOI: 10.1016/bs.adcom.2015.11.002]

[37] C. Kulatunga, K. Bhargava, D. Vimalajeewa, S. Ivanov, “Cooperative in-network computation in energy harvesting device clouds”, Elsevier, Sustainable Computing: Informatics and Systems, vol. 16, pp. 56-65, December 2017.

[38] L. Fan, X. Lei, N. Yang, T. Q. Duong, G. K. Karagiannidis, “Secure Multiple Amplify-and-Forward Relaying With Cochannel Interference”, IEEE Journal of Selected Topics in Signal Processing, vol. 10, no. 8, pp. 1494-1505, December 2016.

[39] M. Haghighat, S. Zonouz, M. Abdel-Mottaleb, "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification", Expert Systems with Applications, vol. 42, no. 11, pp. 7905-7916, November 2015.

[40] C.-z. Gao, Q. Cheng, P. He, W. Susilo, J. Li, “Privacy-Preserving Naive Bayes Classifiers Secure against the Substitution-then-Comparison Attack”, Information Sciences, vol. 444, pp. 72-88, May 2018.

[41] B. Li, Y. Huang, Z. Liu, J. Li, Z. Tian, S.-M. Yiu, “HybridORAM: Practical Oblivious Cloud Storage with constant bandwidth”, Elsevier, Information Sciences, vol. 479, pp. 651-663, April 2019. [DOI: 10.1016/j.ins.2018.02.019]

[42] Y. Kumar, R. Munjal, H. Sharma, “Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures”, IJCSMS International Journal of Computer Science and Management Studies, Vol. 11, Issue: 3, October 2011.

[43] R. Kaur, S. Kinger, "Analysis of Security Algorithms in Cloud Computing", International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 3, pp. 171-176, March 2014.

[44] J. Li, X. Chen, S. S. M. Chow, Q. Huang, D. S. Wong, Z. Liu, “Multi-authority fine-grained access control with accountability and its application in cloud”, Journal of Network and Computer Applications, vol. 112, pp. 89-96, June 2018.

[45] J. Li, Z. Liu, X. Chen, X. Tan, D. S. Wong, “L-EncDB: A Lightweight Framework for Privacy-Preserving Data Queries in Cloud Computing”, Elsevier, Knowledge-based Systems, vol. 79, pp: 18-26, May 2015. [DOI: 10.1016/j.knosys.2014.04.010]

[46] M. Shen, X. Tang, L. Zhu, X. Du, M. Guizani, “Privacy-Preserving Support Vector Machine Training Over Blockchain-Based Encrypted IoT Data in Smart Cities”, IEEE Internet of Things Journal, vol. 6, issue: 5, pp. 7702-7712, October 2019. [DOI: 10.1109/JIOT.2019.2901840]

[47] Q. Xia, E. B. Sifah, K. O.-B. O. Agyekum, H. Xia, K. N. Acheampong, A. Smahi, J. Gao, X. Du, M. Guizani, “Secured Fine-Grained Selective Access to Outsourced Cloud Data in IoT Environments”, IEEE Internet of Things Journal, vol. 6, issue: 6, pp. 10749-10762, December 2019. [DOI: 10.1109/JIOT.2019.2941638]

[48] D. Liu, Z. Yan, W. Ding, M. Atiquzzaman, “A Survey on Secure Data Analytics in Edge Computing”, IEEE Internet of Things Journal, vol. 6, issue: 3, pp. 4946 – 4967, June 2019. [DOI: 10.1109/JIOT.2019.2897619]

[49] C. Stergiou, K. E. Psannis, “Recent advances delivered in Mobile Cloud Computing’s Security and Management challenges”, IGI Global, Modern Principles, Practices, and Algorithms for Cloud Security, 2019.

[50] C. Esposito, M. Ficco, A. Castiglione, F. Palmieri, H. Lu, “Loss-Tolerant Event Communications Within Industrial Internet of Things by Leveraging on Game Theoretic Intelligence”, IEEE Internet of Things Journal, vol. 5, issue: 3, pp. 1679-1689, June 2018. [DOI: 10.1109/JIOT.2017.2782264]

[51] M. Abbasi, A. Shokrollahi, M. R. Khosravi, V. G. Menon, “High-performance flow classification using hybrid clusters in software defined mobile edge computing”, Elsevier, Computer Communications, vol. 160, pp. 643-660, July 2020. [DOI: 10.1016/j.comcom.2020.07.002]

[52] Z. Luo, M. L. Wang, Z. Lin, L. Huang, X. Du, M. Guizani, “Energy-Efficient Caching for Mobile Edge Computing in 5G Networks”, Applied Sciences, vol. 7, issue 6, pp. 1-13, June 2017.

[53] G. Kokkonis, K. E. Psannis, M.Roumeliotis, D.Schonfeld, “Real-time wireless multisensory smart surveillance with 3D-HEVC streams for internet-of-things (IoT)”, Springer, Jounral of Supercompuitng, vol. 73, issue: 3, pp. 1044-1062, March 2017.

[54] G. Kokkonis, K. E. Psannis, M. Roumeliotis, Y. Ishibashi, “Efficient algorithm for transferring a real-time HEVC stream with haptic data through the internet”, Springer, Journal of Real-Time Image Processing, vol. 12, issue: 2, August 2016.

[55] A. V. Sajitha, Dr. A. C. Subhajini, “Analysis of Cloud Sim Toolkit for Implementing Energy Efficient Green Cloud Data Centers”, International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 6, issue IV, pp. 4613-4624, April 2018.

[56] M. R. Rahimi, J. Ren, C. H. Liu, A. V. Vasilakos, N. Venkatasubramanian, “Mobile Cloud Computing: A survey, State of Art and Future Directions”, Mobile Networks and Applications, Volume 19, Issue 2, pp. 133-143, April 2014.

[57] Md Whaiduzzaman, M. N. Haque, Md R. K. Chowdhury, A. Gani “A Study on Strategic Provision of Cloud Computing Services”, The Scientific World Journal, pp. 1-8, June 2014.

[58] S. K. Garg, S. Versteeg, R. Buyya, “A framework for ranking of cloud computing services”, Elsevier, Future Generation Computer Systems, vol. 29, issue: 4, pp. 1012–1023, June 2013. [DOI: 10.1016/j.future.2012.06.006]

[59] G. Singh, S. Kinger, “Integrating AES, DES, and 3-DES Encryption Algorithms for Enhanced Data Security”, International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July 2013.

[60] C. Stergiou, K. E. Psannis, “Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey”, Wiley, International Journal of Network Management, pp. 1-12, May 2016. [DOI:10.1002/nem.1930]

[61] P. Viswanathan, “Cloud Computing – Is it Really All That Beneficial?”, abouttech, 7/7/2012. [Online]. Available: http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm. [Accessed 24/5/2017].

[62] F. Pfarr, T. Buckel, A. Winkelmann, "Cloud Computing Data Protection – A Literature Review and Analysis", in Proceedings of 47th Hawaii International Conference on System Sciences, pp. 5018-5027, 6-9 January 2014, Waikoloa, HI, USA.

[63] E. Almrot, S. Andersson, "A study of the advantages & dis-advantages of mobile cloud computing versus native environment", Digitala Vetenskapliga Arkivet, Bachelor Thesis in Software Engineering, Blekinge Institute of Technology, Karlskrona, May 2013.

[64] Blog: Follow what's happening at Get Cloud Services, "Mobile Cloud Computing – Pros and Cons", GetCloud Services, 23/12/2014. [Online]. Available: https://www.getcloudservices.com/blog/mobile-cloud-computing-pros-and-cons/. [Accessed 24/12/2017].

[65] J. Li, L. Huang, Y. Zhou, S. He, Z. Ming, "Computation partitioning for mobile cloud computing in big data environment", IEEE Transactions on Industrial Informatics, Vol. 11, January 2017.

[66] E. Shi, Y. Niu, M. Jakobsoon, R. Chow, "Implicit Authentica-tion through Learning User Behavior", ACM, in Proceedings of ISC'10 13th International Conference on Information Security, pp. 99-113, 25-28 October 2010, Boca Raton, FL, USA.

[67] A. Botta, W. de Donato, V. Persico, A. Pescape, "Integration of Cloud Computing and Internet of Things: a Survey", Journal of Future Generation Computer Systems, vol. 56, pp. 1-54, March 2016. [DOI: 10.1016/j.future.2015.09.021]

[68] T. Bhattasali, R. Chaki, N. Chaki, "Secure and trusted cloud of things", In Proceedings of INDICON Annual IEEE India Conference, 13-15 December 2013, Mumbai, India. [DOI: 10.1109/INDCON.2013.6725878]

[69] Y. Simmhan, A. G. Kumbhare, B. Cao, V. Prasanna, "An analysis of security and privacy issues in smart grid software architectures on clouds", In Proceedings of IEEE 4th International Conference on Cloud Computing, 4-9 July 2011, pp. 582–589, Washington, DC, USA.

[70] N. Grozev, R. Buyya, "Inter-cloud architectures and application brokering: taxonomy and survey", Wiley Online Library, Software: Practice and Experience, vol. 44, issue: 3, pp. 369–390. March 2014.

[71] B. B. P. Rao, P. Saluia, N. Sharma, A. Mittal, S. V. Sharma, "Cloud computing for Internet of Things & sensing based applications", In Proceedings of IEEE 6th International Conference on Sensing Technology (ICST 2012), pp. 374–380, 18-21 December 2012, Kolkata, India.

[72] W. He, G. Yan, G., L. D. Xu, “Developing vehicular data cloud services in the IoT environment”, IEEE Transactions on Industrial Informatics, vol. 10, issue: 2, pp. 1587–1595, May 2014. [DOI: 10.1109/TII.2014.2299233]

[73] C. Dobre, F. Xhafa, “Intelligent services for big data science”, Elsevier, Future Generation Computer Systems, vol. 37, pp. 267–281, July 2014.

[74] G. Aceto, A. Botta, W. de Donato, A. Pescape, “Cloud monitoring: A survey”, Elsevier, Computer Networks, vol. 57, issue: 9, pp. 2093–2115, June 2013.

[75] P. Mell, T. Grance, “The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology”, National Institute of Standards and Technology Special Publication, pp. 1-7, September 2011.

[76] G. Skourletopoulos, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, J. N. Sahalos, “An Evaluation of Cloud-Based Mobile Services with Limited Capacity: A Linear Approach”, Springer, Soft Computing, vol. 21, issue: 16, pp. 4523-4530, August 2017. [DOI: 10.1007/s00500-016-2083-4]

[77] K. E. Psannis, C. Stergiou, B. B. Gupta, “Advanced Media-based Smart Big Data on Intelligent Cloud Systems”, IEEE Transaction on Sustainable Computing, vol. 4, Issue: 1, pp. 77-87, January-March 2019.

[78] C. Stergiou, K. E. Psannis, A. P. Plageras, T. Xifilidis, B. B. Gupta, “Security and Privacy of Big Data for Social Networking Services in Cloud”, in Proceedings of IEEE conference on Computer Communications (IEEE INFOCOM 2018), 15-20 April 2018, Honolulu, HI, USA.

[79] J. M. Batalla, G. Mastorakis, C. X. Mavromoustakis, J. Zurek, “On cohabitating networking technologies with common wireless access for Home Automation Systems purposes” IEEE Wireless Communications, vol. 23, issue: 5, pp. 76-83, October 2016. [DOI: 10.1109/MWC.2016.7721745]

[80] C. Stergiou, K. E. Psannis, B. Gupta, Y. Ishibashi, “Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT”, Elsevier, Sustainable Computing, Informatics and Systems, vol. 19, pp. 174-184, September 2018.

[81] X. Yao, C. Huang, L. Sun, “Two-Stream Federated Learning: Reduce the Communication Costs”, in Proceedings of 2018 IEEE Visual Communications and Image Processing (VCIP), 9-12 December 2018, Taichung, Taiwan, Taiwan. [DOI: 10.1109/VCIP.2018.8698609]

[82] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, M. Jirstrand, “A Performance Evaluation of Federated Learning Algorithms”, in Proceedings of DIDL '18:

Proceedings of the Second Workshop on Distributed Infrastructures for Deep Learning, December 2018, pp. 1-8, Middleware '18: 19th International Middleware Conference Rennes France. [DOI: 10.1145/3286490.3286559]

[83] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, JMLR: W&CP, volume 54, 20-22 April 2017, Fort Lauderdale, Florida, USA. [arXiv:1602.05629]

[84] S. Thakur, J. G. Breslin, "A Robust Reputation Management Mechanism in the Federated Cloud", IEEE Transactions on Cloud Computing, vol. 7, issue: 3, pp. 625-637, July-September 2019. [DOI: 10.1109/TCC.2017.2689020]

[85] R. Shokri, V. Shmatikov, "Privacy-preserving deep learning", in Proceedings of 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), 30 September – 2 October 2015, Allerton Park and Conference Center, USA.

Chapter 2

Algorithms for Big Data in Advanced Communication Systems and Cloud Computing

C. Stergiou, K. E. Psannis, "Algorithms for Big Data in Advanced Communication Systems and Cloud Computing", in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece. [DOI: 10.1109/CBI.2017.28]

This paper achieves a Commendation for Doctoral Student Work in 19th IEEE Conference on Business Informatics (CBI2017)

2.1 Introduction

Communication Systems are becoming significant in many areas of modern everyday life. Through this field, several technologies grow and contribute to the improvement of people's everyday life. Big Data (BD) appears as a technology created and developed through communication systems. Big Data refers to the large-scale amounts of data used, transferred, and managed through a network. Due to data usage and even large-scale amounts of data, the use of storage space without restrictions on its use becomes necessary. This storage space is provided through a technology called Cloud Computing (CC). This technology mentions a substructure in which data storage and data processing occur in real-time outside of the user's device. This work surveys BD and CC and their basic features, with a focus on the security and privacy issues of both technologies. In addition to this, we will try to combine the functionality of BD and CC to examine the frequent features, and also to discover the benefits related to the security issues of their integration.

2.2 Background Review

2.2.1 Big Data

BD is a more complicated world because the scale is much larger. The information is usually spread out over a number of servers, and the work of compiling the data must be coordinated among them. In the past, the work was largely delegated to the database software, which would use its magical JOIN [11] mechanism to compile tables, after add up the columns before handing off the rectangle of data to the reporting software that would paginate it. This was often harder than it sounds. Database programmers can tell you the stories about complicated JOIN commands that would lock up their database for hours as it tried to produce a report for the boss who wanted his columns just so [8] [12].

2.2.2 Big Data Characteristics

The three Vs of big data (figure 2.1), which are volume, variety, and velocity, constitute a comprehensive definition, and they bust the myth that big data is only about data volume.

Big Data Volume

The Data volume measures the amount of data available to an organization, which does not necessarily have to own all of it as long as it can access it [8] [12] [13].

Big Data Velocity

The Data velocity measures the speed of data creation, streaming, and aggregation. Data velocity management is much more than a bandwidth issue; it is also an ingest issue (extract-transform-load) [8] [12] [13].

Big Data Variety

The Data variety is a measure of the richness of the data representation – text, images video, audio, etc. Incompatible data formats, non-aligned data structures, and inconsistent data semantics represents significant challenges that can lead to analytic sprawl [8] [12] [13].

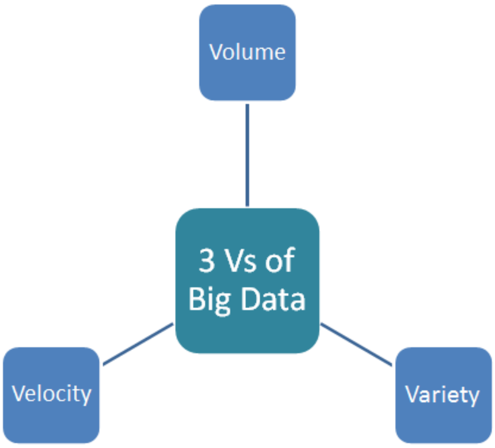


Figure 2.1: The Three Vs of Big Data.

2.2.3 Big Data Security & Privacy

New challenges and standards developed and created in data security issues through the development and the use of BD technology. This creates a growing need for further research on security technologies in order to be able to handle the huge amount of data and to ensure effective. Technologies for securing data are slow when applied to huge amounts of data.

2.2.4 Cloud Computing

Cloud computing provides computing, storage, services, and applications over the Internet (figure 2.2). In general, to render smartphones energy efficient and computationally capable, major changes to the hardware and software level required. This entails the cooperation of developers and manufacturers. [14] [15]. The technology of Cloud computing is the outcome of interdisciplinary approaches combining mobile computing with cloud computing. Thus, this transdisciplinary domain is also referred as mobile cloud computing [8] [14] [16] [17].

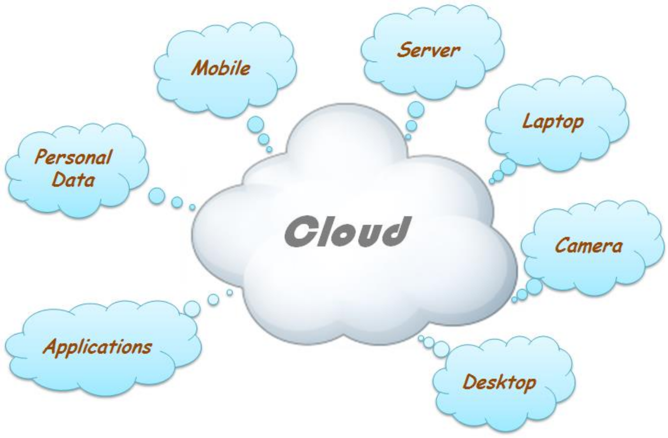


Figure 2.2: Cloud Computing.

2.2.5 Cloud Computing Features

As all technologies, so the CC technology has some characteristics which determine its function. These features are analyzed and outlined subsequently.

Storage over Internet

Storage over Internet can be defined as a technology framework that uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks to link servers and storage devices, and to facilitate storage solution deployment. The Storage over Internet technology is also known as Storage over Internet Protocol (SoIP) technology [18] [19] [20].

Service over Internet

The main objective of the Service over Internet is to be committed to help customers all over the world with the aim to transform aspirations into achievements by harnessing the Internet’s efficiency, speed and ubiquity [18] [19] [21].

Applications over Internet

The programs which can be written to do the job of a current manual task, or virtually anything, and which perform their job on the server (cloud server) via an internet connection rather than the traditional model of a program that has to be installed and run on a local computer are the Cloud Applications, or as a scientific definition Applications over Internet [18] [19] [20].

Energy Efficiency

As a definition, the Energy Efficiency is a way of managing and restraining the growth in energy consumption. By delivering more services for the same energy input or for the same services for less energy input may be something more energy efficient [18] [19] [20].

<u>Computationally Capable</u>

The services of computational clouds are leveraging the computationally intensive and ubiquitous mobile applications which have been enabled by the technology of Mobile Cloud Computing [6] [17] [18] [19].

2.2.6 Cloud Computing Security Issues

CC and storage solutions provide users and enterprises with various capabilities to store and process their data in third-party data centers [22] [23]. Organizations use the Cloud in a variety of different service models (SaaS, PaaS, and IaaS) and deployment models (Private, Public, Hybrid, and Community) [15] [24] [25]. There are a number of security concerns associated with cloud computing. These issues fall into two broad categories: security issues faced by cloud providers (organizations providing software, platform, or infrastructure-as-a-service via the cloud) and security issues faced by their customers (companies or organizations who host applications or store data on the cloud) [26] [27]. The responsibility is shared, however. The provider must ensure that their infrastructure is secure and that their clients' data and applications are protected while the user must take measures to fortify their application and use strong passwords and authentication measures [8] [9].

2.3 Research Objectives

There is a survey of BD and CC technologies and their basic characteristics, with a focus on the security and privacy issues of both technologies. Moreover, I will try to combine the functionality of the two aforementioned technologies (i.e Big Data and Cloud Computing) in order to examine the common features, and also to discover the benefits related in security issues of their integration. The main goal of this paper is to try to combine the functionality of the BD and CC technologies in order to examine the common features, and also to discover the benefits related in security issues of their integration. This could be take place by the presentation of a new method of an algorithm that can be used for the purpose of improving Cloud Computing's security through the use of algorithms that can provide more privacy in the data related to Big Data technology. Furthermore, we survey the security challenges of the integration of those technologies. This can be the field of future research on the integration of those two technologies, and why not to have a huge improvement of their security and privacy issues in order to have a better use of them.

2.4 Research Approach & Methods

The purpose of this research proposal is to study the technologies for the "*large-scale data*" (big data) and "*cloud computing*" in order to analyze and manage the telecommunication systems.

This study will rely on the following items in order to examine all the existing and the future data in order to gather and produce all the necessary tools with the aim to come out:

1. Action Research
2. Development and implementation of the Unified Modeling Language (UML)
3. Software engineering
4. Implementation and model development
5. Usability evaluation

According to the aforementioned data, the survey will rely on the study of existing literature and the use of existing research on the fields I will study, in order to use all the necessary information. The next stage will be the recovery process of all the material will be collected and organized in the different steps. Based on the current situation regarding the scope of my research will develop simulation models in order to study and achieve my proposal. Finally, there will be a verification of the data that will be produced during the implementation of the model and the usability of the proposed model will be studied as well as the further improvement and evaluation.

2.5 Current & Expected Contributions

As already mentioned, with this work I will try to find a better security algorithm model for the BD in Cloud environments. For this purpose several existing security algorithms would be studied.

Table 2.1 lists a sample of the most popular security algorithms that would be studied in order to produce a new algorithm model. As regards the Table 2.1 we could conclude that even the most efficient algorithms give an encryption rate of 64.3MB/s. The information of those algorithms have been taken through related works [8] [18] [19] [21] [31].

Algorithm	*Key length*	*Megabytes processed*	*Block size*	*Rounds*
Blowfish	32-448 bits	256	64 bits	16
DES	56 bits	128	64 bits	16
3-DES	56, 112 or 168 bits	128	64 bits	48
AES	128, 192 or 256 bits	256	128 bits	10, 12 or 14
RSA	1025 – 4096 bits	300	512 bits	1

Table 2.1: Encryption Rates of popular Security Algorithms.

Through the table 2.1 we can conclude that in the sector of BD technology, in which the need of large amounts of data needs to be transferred we could observe an important bottle neck for encryption like huge amounts of data. This is harmful to the nature of BD that have real time processing and outcomes.

Subsequently, a correlation of characteristics of BD and CC can be made.

Big Data Features	***Volume***	***Velocity***	***Variety***
Cloud Computing Features			
Storage over Internet		X	
Service over Internet	X		X
Applications over Internet	X	X	X
Energy Efficiency	X	X	
Computational Capable		X	X

Table 2.2: Correlation of BD and CC characteristics.

By the table 2.2 can be exhibited the key characteristics of the two technologies which have been studied and used for the experimental proposal. Count on the study conducted, the key feature of BD technology which contributes more with the characteristics of CC technology is Velocity. Velocity contributes four from the five key characteristics of CC. Also, another thing that we can observe from table 2.2 is that the characteristic Applications over Internet contributed from all the key features of BD.

Moreover, a big part of this work would relay on the related works that have been made previously. Table 2.3 lists the findings and the concepts associated with problems and respective solutions indicating in a previous works. In table 2.3 the former work list in ascending chronological order starting from 2010 until today.

Author	Problems	Solutions
H. Takabi et al **[32]**	• Specific characteristics worsen security & privacy challenges of Cloud Computing.	• Examines the possibilities of offering a trustworthy CC environment.
H. T. Dinh et al **[33]**	• Detonating growth of mobile applications & resurgent of CC concept is considered advancement in mobile services.	• A survey of MCC, with focus on its definition, architecture & applications.
N. Fernando et al **[34]**	• Intrinsic problems (e.g. resource scarcity, frequent disconnections) hinter the usage of mobile computing in its full scale.	• Categorizes the major issues in MCC & discusses different methods to solve these issues. • Careful examination of problems which have not yet been addressed & put forward ideas for future research.
Sachdev & M. Bhansali **[35]**	• The bigger the number of cloud users the most frequent the malicious activity in the cloud. • Highly safe and persistent services needed.	• A data encryption model which protects the privacy and security of the data before they are uploaded in the cloud.
M. Ali et al **[p13-36]**	• Third-party cloud services have more deficiencies and more vulnerable to security threads. • Sharing the users' data outside the administrative control.	• Examines and shortly analyzes both internal and external security problems in the Mobile Cloud Computing.
S. Bhavani et al **[37]**	• Load balancing is one of the cloud's issues. • A reduction in the response time and optimization of the resource utilization can be achieved balancing the load.	• The best algorithm for balancing the load is Ant Colony Optimization.
S. Sathya & R. Avinash **[38]**	• How people adopt cloud as Cloud Technologies Mature.	• An explanation of how BD and cloud responds for user's demand as a compelling combination.
S. Rallapalli et al **[39]**	• The healthcare organizations face the critical challenge to analyze big data. • Large amounts of data cannot be processed through conventional systems.	• Hadoop: An application which could prepare huge amounts of data in distributed environment could be deployed on cloud environment to prepare the big amount of healthcare data.
O. Awodele et al **[40]**	• Security challenges are the most serious in cloud & big data services. • Issues of service level agreement.	• Shipping disk drives to cloud computing. • Use of Data mining techniques. • Use of Access control techniques.
N. R. Vajjhala & E. Ramollari **[41]**	• Contemporary methods in the field of BD using cloud resources. • How the SMEs can take advantage of these technological trends.	• Cloud computing offers an alternative to SMEs shifting the burden of providing and maintaining expensive infrastructure to cloud service providers.
P. Zhou et al **[42]**	• The increased usage of social media has created a new period, that of the BD. • Privacy of users' contexts & video service sellers' repositories, that are remarkably sensitive & of important commercial value.	• An innovative "geometric differentially private" scheme, that could minimize the performance loss.
A. A. Gnana Singh et al **[43]**	• Promote the research and development activities in the field of BD and CC.	• A method for storing the data on cloud using the cloudsim package.

Table 2.3: Mapping problems against referenced solutions.

2.6 Chapter Summary

The CC technology provides many possibilities, but in addition to this places quite a lot of restrictions as well. This technology mentions to an infrastructure where both the data storage and processing occur outside of the user's device. In this work, we survey BD and CC technology and their basic characteristics, with a focus on the security and privacy issues of both technologies. Moreover, we have tried to combine the functionality of the two aforementioned technologies (i.e BD & CC) with the aim to examine the frequent characteristics, and moreover to discover the benefits related in security issues of their integration.

The main goal of this work is to find novel ways to achieve a better integration of BD and CC, with focus on security algorithms and all the challenges that the two aforementioned technologies faced on security level. This can be the field of future research on the integration of those two technologies. Regarding the rapid development of both technologies the security issue must be solved or reduced to a minimum in order to have a better integration model. These security challenges that surveyed in this paper could be the sector for further research as a case study, with the goal of minimizing them.

2.7 Chapter References

[6] C. Stergiou, K. E. Psannis, "Mobile Cloud Computing in 4G Networks (LTE)", (2015), in Proceedings of 2nd Student Conference of Applied Informatics, 2 December 2015, Thessaloniki, Greece.

[8] C. Stergiou, K. E. Psannis, "Efficient and Secure Big Data delivery in Cloud Computing", Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017. [DOI:10.1007/s11042-017-4590-4]

[9] D. Tomtsis, S. Kontogiannis, G. Kokkonis, I. Kazanidis, S. Valsamidis, "Proposed cloud infrastructure of wearable and ubiquitous medical services", in Proceedings of 5th Fifth International Conference on Digital Information Processing and Communications (ICDIPC 2015), pp. 213-218, 7-9 October 2015, Sierre, Switzerland.

[11] Y. Kryftis, G. Mastorakis, C. Mavromoustakis, J. Mongay Batalla, E. Pallis, G. Kormentzas, "Efficient Entertainment Services Provision over a Novel Network Architecture", IEEE Wireless Communications, vol. 23, issue: 1, pp. 14-21, March 2016. [DOI: 10.1109/MWC.2016.7422401]

[12] C. Buckler, "Understanding JOINs in MySQL and Other Relational Databases", sitepoint, 19/5/2011. [Online]. Available: https://www.sitepoint.com/understanding-sql-joins-mysql-database/. [Accessed 21/5/2016].

[13] P. Russom, “Big Data Analytics”, TDWI RESEARCH, TDWI Best Practices Report, Fourth Quarter 2011, USA.

[14] D. Huang, “Mobile cloud computing,” IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, vol. 6, no. 10, pp. 27–31, 2011.

[15] C. Stergiou, K. E. Psannis, B.-G. Kim, B. B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI:10.1016/j.future.2016.11.031]

[16] M. A. Alsmirat, Y. Jararweh, I. Obidat, B. B. Gupta, “Internet of Surveillance: A Cloud supported Large Scale Wireless Surveillance System”, Springer, The Journal of Supercomputing, vol. 73, pp. 973-992, September 2016. [DOI: 10.1007/s11227-016-1857-x]

[17] C. Stergiou, “Technologies of Internet of Things and Mobile Cloud Computing”, Bachelor Dissertation, Technology Management, Information Technology, University of Macedonia, June 2016.

[18] Md Whaiduzzaman, M. Nazmul Haque, Md R. Karim Chowdhury, A. Gani “A Study on Strategic Provision of Cloud Computing Services”, Hindawi, The Scientific World Journal, vol. 2014, pp. 1-8, Article ID 894362, June 2014. [DOI: 10.1155/2014/894362]

[19] S. K. Garg, S. Versteeg, R. Buyya, “A framework for ranking of cloud computing services”, Elsevier, Future Generation Computer Systems, Vol. 29, issue: 4, pp. 1012–1023, June 2013. [DOI: 10.1016/j.future.2012.06.006]

[20] G. Skourletopoulos, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, J. N. Sahalos, “An Evaluation of Cloud-Based Mobile Services with Limited Capacity: A Linear Approach”, Springer, Soft Computing, vol. 21, issue: 16, pp. 4523-4530, August 2017. [DOI: 10.1007/s00500-016-2083-4]

[21] L. Borovick, R. L. Villars, “The Critical Role of the Network in Big Data Applications”, IDC Analyze the Future, White Paper, Sponsored by: Cisco Systems pp. 1-12, April 2012.

[22] M. Haghighat, S. Zonouz, M. Abdel-Mottaleb, “CloudID: Trustworthy cloud-based and cross-enterprise biometric identification”, Elsevier, Expert Systems with Applications, vol. 42, Issue: 21, pp. 7905-7916, November 2015. [DOI: 0.1016/j.eswa.2015.06.025]

[23] K. Bhushan, B. B. Gupta, “Security Challenges in Cloud Computing: State-of-art”, InderScience Publisher, International Journal of Big Data Intelligence (IJBDI), vol. 4, No. 2, March 2017.

[24] M. K. Srinivasan, K. Sarukesi, P. Rodrigues, M. S. Manoj, P. Revathy, "State-of-the-art cloud computing security taxonomies: a classification of security challenges in the present cloud computing environment", in Proceedings of the International Conference on Advances in Computing, Communications and Informatics ICACCI '12, pp. 470-476, August 2012. [DOI: 10.1145/2345396.2345474]

[25] Z. Gou, S. Yamaguchi, B. B. Gupta, "Analysis of Various Security Issues and Challenges in Cloud Computing Environment: A Survey," IGI Global, Book Chapter, Handbook of Research on Modern Cryptographic Solutions for Computer and Cyber Security, pages: 27, USA, 2016. [DOI: 10.4018/978-1-5225-0105-3.ch017]

[26] Y. Mamoon, "Swamp Computing" a.k.a. Cloud Computing", WEB Security Journal, 28/12/2009. [Online]. Available: http://security.sys-con.com/node/1231725. [Accessed 27/07/2016].

[27] R. Shaikha, Dr. M. Sasikumar, "Data Classification for achieving Security in cloud computing", Elsevier, Procedia Computer Science, vol. 45, pp. 493-498, 2015. [DOI: 10.1016/j.procs.2015.03.087]

[31] O. Badve, B. B. Gupta BB, S. Gupta, "Reviewing the security features in contemporary security policies and models for multiple platforms", IGI Global, Book Chapter, Handbook of research on Modern Cryptographic Solutions for Computer and Cyber Security, pages: 26, USA, 2016. [DOI: DOI: 10.4018/978-1-5225-0105-3.ch020]

[32] H. Takabi, J. B. D. Joshi, "Security and Privacy Challenges in Cloud Computing Environments", IEEE Security & Privacy, vol. 8, Issue: 6, pp. 24-31, November-December 2010. [DOI: 0.1109/MSP.2010.186]

[33] H. T. Dinh, C. Lee, D. Niyato, P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches", Wiley, Wireless Communications & Mobile Computing, vol. 13, Issue: 18, pp. 1587-1611, October 2011. [DOI: 10.1002/wcm.1203]

[34] N. Fernando, S. W. Loke, W. Rahayu, "Mobile cloud computing: A survey", Elsevier, Future Generation Computer Systems, vol. 29, Issue: 1, pp. 84-106, January 2013. [DOI: 10.1016/j.future.2012.05.023]

[35] A. Sachdev, M. Bhansali, "Enhancing Cloud Computing Security using AES Algorithm", International Journal of Computer Applications, vol. 9, no. 67, pp. 19-23, April 2013.

[36] M. Ali, S. U. Khan, A. V. Vasilakos, “Security in cloud computing: Opportunities and challenges”, Elsevier, Information Sciences, vol. 305, pp. 357-383, June 2015. [DOI: 10.1016/j.ins.2015.01.025]

[37] Prof. S. Bhavani, A. Hatwal, U. Mittal, “"Study on Cloud Computing and Different Load Balancing Algorithms in Cloud Computing”, International Journal of Emerging Research in Management &Technology, vol. 4, Issue: 5, pp. 331-336, May 2015. [ISSN: 2278-9359]

[38] S. Sathya, R. Avinash, “Big Data and Cloud Computing”, in Proceedings of Rathinam College National Conference, 2015, Pollachi Road, Echanari, India.

[39] S. Rallapallia, R. R. Gondkar, U. P. K. Ketavarapu, “Impact of Processing and Analyzing Healthcare Big Data on Cloud Computing Environment by Implementing Hadoop Cluster”, Elsevier, Procedia Computer Science, vol. 85, pp. 16-22, 2016. [DOI: 10.1016/j.procs.2016.05.171]

[40] A. Oludele, A. Izang, S. O Kuyoro, F. Y. Osisanwo, “Big Data and Cloud Computing Issues”, International Journal of Computer Applications, vol. 12, no. 133, pp. 14-19, January 2016. [DOI: 10.5120/ijca2016907861]

[41] Dr. N. R. Vajjhala, Dr. E. Ramollari, “Big Data using Cloud Computing - Opportunities for Small and Medium-sized Enterprises”, European Journal of Economics and Business Studies, vol. 2, issue: 1, pp. 129-137, January-April 2016. [ISSN 2411-4073]

[42] P. Zhou, Y. Zhou, D. Wu, H. Jin, “Differentially Private Online Learning for Cloud-Based Video Recommendation With Multimedia Big Data in Social Networks”, IEEE Transactions on Multimedia, vol. 18, Issue: 6, pp. 1217-1229, June 2016. [DOI: 10.1109/TMM.2016.2537216]

[43] R. Iqbal, F. Doctor, B. More, “Big Data analytics: Computational intelligence techniques and application areas”, International Journal of Information Management, pp. 1-11, June 2016. [DOI: 10.1016/j.ijinfomgt.2016.05.020]

Chapter 3

Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey

C. Stergiou, K. E. Psannis, "Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey", Wiley, International Journal of Network Management, pp. 1-12, May 2016. [DOI:10.1002/nem.1930]

3.1 Introduction

The Internet of Things is a new technology that is growing rapidly in the field of telecommunications, and especially in the modern field of wireless telecommunications. The main goal of the interaction and cooperation between things and objects sent through the wireless networks is to fulfill the objective set to them as a combined entity. In addition, based on the technology of wireless networks, both the technologies of Mobile Cloud Computing and the Internet of Things develop rapidly. In this paper, I combine the two aforementioned technologies (i.e MCC and IoT) with the technology of Big Data to examine the common features, and to discover which of the MCC and IoT benefits improve the use of Big Data Applications. Finally, I present the contribution of MCC and IoT individually to the technology of Big Data.

3.2 Related Work

In recent years they have been made a couple of researches in order to improve the field of telecommunications. In this field there are some technologies that we study and analyze in this paper. As regard the Mobile Cloud Computing a presentation of what it offers about abundant computing power which can be tapped easily have been studied, and also, what systematically explore of the fundamental research questions when mobile and cloud computing combining [10]. Additionally, a study of these fundamental new capabilities which will enable mobile users to seamlessly utilize the cloud to obtain the resource benefits without incurring delays and jitter and without worrying about energy [10]. Moreover, the Cloud Computing technology is now used in the emerging IT platforms, used as a market-oriented resource allocation by leveraging technologies such as Virtual Machines and the insights on market-based resource management strategies that encompass both customer-driven service management and computational risk management to sustain Service Level Agreement (SLA)-oriented resource allocation. [11]. Therefore, the Mobile Cloud Computing technology start to work with J2ME applications [12]. Thus, as long as the research continues the architecture of Mobile Cloud Computing evolves and now the interactions between mobile applications and cloud services will be decoupled and a depiction of the uncoupling of service access and delivery to mobile applications was provided [13].The Cloud Computing technology in nowadays affects the remote resources on the quality and reliability of augmentation processes and discuss the challenges and opportunities of employing varied cloud-based resources in augmenting mobile devices and in a taxonomy of CMA approaches [15]. Finally, to conclude with the technology of the Mobile Cloud Computing, we realized that with the evolution of this technology now we can run runtime application partitioning on SMD by analyzing additional resources utilization on SMD in the mechanism of runtime application profiling and partitioning [16].

Subsequently, as regard the technology of Big Data, we have study the aspects of scheduling and storage, which are the foundations of modern big data analytics systems and their key principles, and how these principles are realized in widely-deployed systems [17]. Moreover, as the field of mobile telecommunications evolves new technologies have been explored and some of them based on the combination of current technologies. Big Data and the technology of Internet of Things have been combined to work together and they create a Cognitive Oriented IoT Big-data Framework (COIB-framework) along with implementation architecture, IoT big-data layering architecture, and data organization and knowledge exploration subsystem for effective data management and knowledge discovery that is well-suited with the large scale industrial automation applications [18]. Also, these two technologies ease the reuse of algorithms and support scientific discussions by providing a comparison schema and the use cases from different industries [20]. In addition, we have study the six phases of location information flow in the IoT and three areas of privacy controls that may be considered in order to manage those flows and so to be helpful to practitioners and researchers when evaluating the issues involved as the technology advances [21]. Finally, we have study and analyze an integrated Big Data analytical framework for Internet of Things and Smart City application, which contributes three things: (1) we provide an overview of Big Data and Internet of Things technologies including a summary of their relationships, (2) we present a case study in the smart grid domain that illustrates the high level requirements towards such an analytical Big Data framework, and (3) we present an initial version of such a framework mainly addressing the volume and velocity challenge. Finally, they present the findings of extended results from the EU funded project BIG and the German funded project PEC [22]. As a conclusion, we have study the combination of the technologies of Mobile Cloud Computing, Big Data and Internet of Things and we have study some analyzation of existing components and methods of securely integrating big data processing with cloud M2M (Machine to Machine) systems based on Remote Telemetry Units (RTUs) [19].

By studying the technologies of Mobile Cloud Computing and Internet of Things, we realize that they have some features which will be able to assist in improving the functionality of Big Data Applications. As a result, these three technologies can be clearly improved in if we combine their use.

3.2.1 Internet of Things

Internet of Things (IoT) is a new technology in telecommunication fields. The Internet of Things (IoT) is the network of physical objects, devices, vehicles, buildings and other items which are embedded with electronics, software, sensors, and network connectivity, which enables these objects to collect and exchange data [23].

The Internet of Things is a network of devices that transmit, share, and use data from the physical environment to provide services to individuals, corporations, and society. The objects-things function either individually or in connection with other objects or individuals, and have unique IDs (identifiers). Also, the Internet of Things has different applications in health, transport, environment, energy or types of devices: sensors, devices worn / carried (wearable), e.g. watch, glasses, home automation (domotics) [24].

The Internet of Things is the next big step in the field of new technology, but the big difference is found in the way businesses operate. Overall, over the next few years a flare in the number of connected devices, the sites are located, and of course the functions they will perform is expected. We can mention future hospitals as an example: in addition to standalone devices connected, there will be numerous devices that will be connected to patient monitoring stations of the nursing staff.

3.2.1.1. Ten Facts and Predictions about the Internet of Things (IoT)

There are many items that are connected through the Internet, and many economic benefits that can be generated from the analysis of the data flows. Consider the following:

- The total added value of the IoT in all sectors will reach 1.9 trillion dollars by 2020 worldwide, according to Gartner.
- Fifty billion devices will be connected to the Internet by 2020, down by Cisco.
- The market of equipment for remote patient monitoring was doubled from 2007 to 2011, and is expected to double again in 2016.
- Smart grids in the energy sector are expected to double the market information systems of customers from $ 2.5 billion in 2013 to $ 5.5 billion in 2020, according to a study by Navigant Research.
- The widespread use of IoT technologies in the auto industry could save $ 100 billion a year by reducing accidents, according to McKinsey.
- Industry Internet could add 10-15 trillion in world GDP, doubling the US economy, says GE.
- 75% of leaders of the global business world explore the economic potential of the Internet of Things, according to a report by The Economist magazine.
- The UK Government has recently approved £ 45 million in research funding for the technologies of the Internet of Things.
- The cities were to spend $ 41 trillion over the next 20 years on the upgrading of the Internet of Things infrastructure, according to Intel.
- The number of manufacturers involved in activities related to IoT, will reach the number of 1.7 million worldwide by the end of 2014, according to ABI Research.

3.2.1.2. Internet of Things: Take advantage of the data

What does it mean when the devices and sensors are networked together and communicate with each other? How can the Internet of Things affect our daily life? GPS systems, alarm systems and thermostats, all send and receive constant feeds to monitor and automate activities in our daily lives [20]. And the not so obvious: Mosaic, cups, clothes and other everyday objects can also join network to send and receive data over the Internet.

Opportunities where the streaming data will create new markets in order to inspire positive change or to enhance existing services are examined by businesses [26]. Bellow there are some examples of sectors that are at the heart of these developments:

- Smart solutions in the bucket of transport, achieving a reduction of traffic on the roads, reducing fuel consumption, set priorities in vehicle repair programs, and save lives.
- Smart power grids incorporating more renewables, improve system reliability and reduce the charges consumers, thus providing cheaper electricity.
- Remote monitoring of patients, providing easy access to health care, improving the quality of services, increasing the number of people served, and saving money.
- Sensors in homes and airports, or even in your shoes or doors, improve safety by sending signals when left unused for a certain period of time or when used in the wrong time.
- Engine monitoring sensors that detect and predict maintenance issues, inventory replenishment, and even define priorities in scheduling maintenance work, repairs and regional operations.

3.2.2 Mobile Cloud Computing

Cloud computing provides computing, storage, services, and applications over the Internet. In general, to render smartphones energy efficient and computationally capable, major changes to the hardware and software levels are required. This entails the cooperation of developers and manufacturers. [27]. Mobile cloud computing is defined as an integration of cloud computing technology with mobile devices in order to make the mobile devices resource-full in terms of computational power, memory, storage, energy, and context awareness. The technology of Mobile Cloud computing is the outcome of interdisciplinary approaches combining mobile computing with cloud computing. Thus, this transdisciplinary domain is also referred as mobile cloud computing [27].

There are two perspectives in which the term Mobile Cloud refers: a) infrastructure based, and b) ad-hoc mobile cloud. In the infrastructure based mobile cloud, the hardware infrastructure remains static and also provides services to the mobile users. Nevertheless, there are several applications which utilize cloud resources, but the usage is limited to only storage and application-specific services such as Apple's Siri (voice based personal assistant) and iCloud storage service.

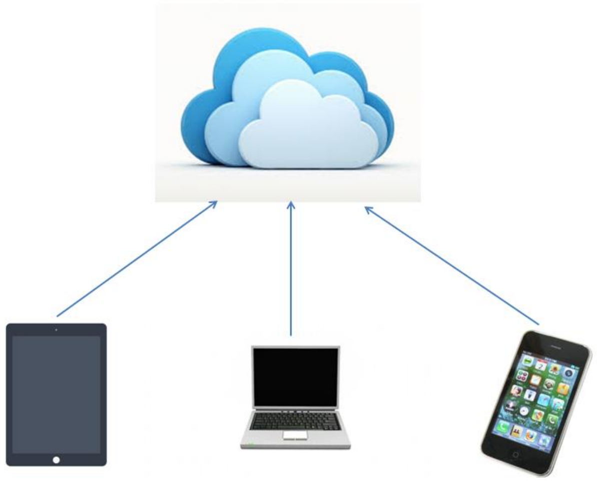


Figure 3.1: Mobile Cloud Computing Technology.

3.2.2.1. Mobile Cloud Computing Trade Offs

Mobile Cloud Computing has some disadvantages-limitations which should be eliminated over the years in order to achieve a better and more ideal use. Some businesses, especially the smaller ones, need to be aware of these limitations before going in for this technology.

A. Security

A major issue of the Mobile Cloud Computing is the security issue. Before someone adopts this technology, they should know that all the company's sensitive information would be released to a third-party cloud service provider. This could potentially put the company to great risk. Hence, someone must be absolutely sure that they would choose the most reliable service provider, who will keep the given information given completely safe [28].

B. Connectivity

Internet connection is critical to Mobile Cloud Computing. Thus, the user should be certain that there is a good result before opting for these services. Since owing a mobile device which is connected to the internet has become the norm in the wireless world of today, Mobile Cloud Computing has a very large potential user base [29].

C. Performance

Another major concern of the Mobile Cloud Computing relates to its performance. Some users feel performance is not as good as with native applications. Thus, checking with one service provider and understanding their track record is advisable [30].

D. Latency (Delay)

In mobile cloud computing, latency is defined as the time involved in offloading the computation and getting back the results from the nearby infrastructure or cloud (sometimes referred as turnaround time). The latency depends on multiple factors such as offloaded code size, data input size, location of the required data, offloading scheme and granularity, network bandwidth, execution delay, and resultant data size.

E. Privacy

Data privacy is important and is one of the main bottlenecks that restrict consumers from adopting mobile cloud computing. The users' data stored in the cloud may include emails, tax reports, personal images, salary and health reports etc, and may contain sensitive information. Therefore, the consumers cannot afford any privacy leakage as it may lead to financial loss and legal issues [31]. The European Union has passed some laws [32] for the handling of data, according to which the data storage servers must reside in the countries in order to provide sufficient protection. Moreover, in some cases the data storage location must be known. However, this is not always possible in a cloud environment due to the absence of standards, data privacy, and cloud security [33]. Therefore, to gain consumers trust in the mobile cloud, the application models must support application development with privacy protection and implicit authentication mechanisms [34].

3.2.3 Big Data Applications

More than $15 billion on software firms that specialize in data management and analytics have been spend by AG, Oracle Corporation, IBM, Microsoft, SAP, EMC, HP and Dell. These companies increased their demand for information management specialists on the provided software. In the year 2010, this industry was worth more than $100 billion and was growing at almost 10% in a year, that is about twice as fast as the software business as a whole [35].

The use of data-intensive technologies by the developed economies has increased. Nowadays, there are up to 4 billion mobile phone subscriptions around the world, and approximately 2 billion people have access to the internet [35]. From 1990 to 2005, up to 1 billion people worldwide joined the middle class, and as a consequence more people are characterized as literate, which in turn leads to information development. The predictions of the world's effective capacity to exchange information through telecommunication networks put the amount of internet traffic at 667 exabytes annually by 2014 [35] [36].

There is a large number of examples for the use of Big Data technology in the public service such as : 1) The Joining up data, which is a local authority blended data about services, such as road gritting rotas, with services for people at risk, such as "meals on wheels". The connection of data allowed the local authority to avoid any weather related delay. 2) The Data on prescription drugs. By connecting the origin, the location and the time of each prescription, a research unit was able to exemplify the considerable delay between the releases of any given drug. A specific example is the UK-wide adaptation of the National Institute for Health and Care Excellence guidelines.

3.2.3.1. Predictable & Efficient

Big Data Applications significantly increase the amount of real-time and workload-intensive transactions through the massive amounts of diverse data transferred. The supporting network which connects the hyperscale server architectures, consisting of thousands of nodes which in turn contain several processors, must be robust enough to ensure this data could move quickly and efficiently

The appropriate line rate of performance furthers the network efficiently. One of the essential tasks needed to achieve network efficiency is to rightsize switch capacity. The typical network configurations for the Big Data are likely to require 1GB access layer switch capacity in the current environment. For the upcoming twelve to eighteen months, 10GbE server connectivity would become more common, as the cost-performance ratio becomes more efficient, and it might need to upgrade aggregation switch capacity to 40GbE or even 100GbE by some organizations [37].

3.2.3.2. Holistic Network

When it comes to the optimized network performance, it is necessary that it takes place within the Big Data domain as well as in the connection with the more traditional undertaking infrastructure [37].

The benefits to a holistic network approach include the following:

- ✓ Ability to minimize duplicative costs whereby one network can support all workloads.
- ✓ Multitenancy to consolidate and centralize Big Data projects.
- ✓ Ease of network provisioning where sophisticated intelligence is used to manage workloads based on business priorities.
- ✓ Ability to leverage network staffing expertise across the entire datacenter.

3.2.3.3. Network Partitioning

The separation, without adding cost and complexity, was enabled by logical partitioning. In addition, with the use of hard partitioning on the Ethernet switch, various tasks might also need to be isolated. This means that the tasks are completely separated at the data plane level. As an example, the data plane separation would be necessary to comply with regulations and privacy requirements, associated with healthcare applications which might be contain sensitive data [37].

3.2.3.4. Scale Out

The ability of “junior science projects” (Big Data projects that might start small) to “Scale Out” would ensure a seamless transition as projects increase in size and number. Additionally, an equally important issue is that network performance and ease of management remain constant as the cluster scales. The oversubscription should be minimized within the Big Data cluster network because of the demands of machine-to-machine traffic flows [37].

3.2.3.5. Unified Ethernet Fabrics

Through leveraging multiple paths through the network, and continuously determining the most efficient route, Unified Ethernet Fabrics enables full link utilization. The Unified Ethernet Fabrics offer an excellent scalability since the virtual chassis architectures provide access to multiple switches and, at the same time, manage them as a single device. This creates a pool of virtual switching resources and eliminates the need for manual configuration. Also, predictable any-to-any latency and bandwidth for traffic between servers within the Big Data cluster is provided by this design. Finally, a distributed approach to networking, which is much more resilient to failures, was brought by the Unified Ethernet Fabrics [37].

3.3 Contribution of Internet of Things in Big Data Applications

Internet of Things	**Predictable & Efficient** [3.2.3.1]	**Holistic Network** [3.2.3.2]	**Network Partitioning** [3.2.3.3]	**Scale Out** [3.2.3.4]	**Unified Ethernet Fabrics** [3.2.3.5]
Smart solution in the bucket of transport [3.3.1]	**X**		**X**	**X**	
Smart power grids incorporating more renewable [3.3.2]	**X**	**X**			**X**
Remote monitoring of patients [3.3.3]		**X**	**X**		**X**
Sensors in homes and airports [3.3.4]			**X**	**X**	**X**
Engine monitoring sensors that detect & predict maintenance issues [3.3.5]	**X**		**X**		**X**

Table 3.1: Contributions of Internet of Things in Big Data Applications.

Table 3.1 lists the features of the technology of Things, with regard to the convenience it offers. It also presents some of the most powerful features of Big Data technology which relate to the Big Data applications. The purpose of Table 3.1 is to show which of the particular features of the IoT technology pertain to, and improve the specific features of the Big Data applications. As we can observe, Unified Ethernet Fabrics and Network Partitioning are the Big Data application features which are affected more by the features of the IoT technology. In contrast, the Holistic Network and the Scale out are the two features influenced less by the features of IoT technology.

3.3.1. Smart solution in the bucket of transport

Smart solutions in the bucket of transport, achieving a reduction of traffic on the roads, reducing fuel consumption, set priorities in vehicle repair programs, and save lives [26].

3.3.2. Smart power grids incorporating more renewable

Smart power grids incorporating more renewables, improve system reliability and reduce the charges consumers, thus providing cheaper electricity [26].

3.3.3. Remote monitoring of patients

Remote monitoring of patients, providing easy access to health care, improving the quality of services, increasing the number of people served, and saving money [26].

3.3.4. Sensors in homes and airports

Sensors in homes and airports, or even in your shoes or doors, improve safety by sending signals when left unused for a certain period of time or when used in the wrong time [26].

3.3.5. Engine monitoring sensors that detect & predict maintenance issues

Engine monitoring sensors that detect and predict maintenance issues, inventory replenishment, and even define priorities in scheduling maintenance work, repairs and regional operations [26].

3.4 Contribution of Mobile Cloud Computing in Big Data Applications

Mobile Cloud Computing	**Predictable & Efficient** [3.2.3.1]	**Holistic Network** [3.2.3.2]	**Network Partitioning** [3.2.3.3]	**Scale Out** [3.2.3.4]	**Unified Ethernet Fabrics** [3.2.3.5]
Storage over Internet [3.4.1]		**X**	**X**		**X**
Service over Internet [3.4.2]	**X**	**X**	**X**		
Applications over Internet [3.4.3]	**X**		**X**	**X**	
Energy Efficiency [3.4.4]	**X**	**X**	**X**		
Computationally Capable [3.4.5]	**X**	**X**			**X**

Table 3.2: Contributions of Mobile Cloud Computing in Big Data Applications.

Table 3.2, like Table 3.1, lists the features of the Mobile Cloud Computing technology, regarding the convenience this technology offers. It also enumerates some of the powerful features of Big Data technology which pertain to the same Big Data Applications studied in Table 3.1. The purpose of Table 3.2 is to show which of the specific features of the Mobile Cloud Computing technology, relate to, and improve the characteristics of Big Data Applications that we singled out. As we can observe from Table 3.2, Network Partitioning, Holistic Network and Network Partitioning are the Big Data Application features which are influenced more by the features of Mobile Cloud Computing technology. In contrast, Scale out is the feature affected less by the features of Mobile Cloud Computing technology.

3.4.1. Storage over Internet

Storage over Internet can be defined as a technology framework that uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks to link servers and storage devices, and to facilitate storage solution deployment. The Storage over Internet technology is also known as Storage over Internet Protocol (SoIP) technology. With the combination of the best storage and networking industry approaches, SoIP provides high-performance and scalable IP storage solutions [38].

3.4.2. Service over Internet

The main objective of the Service over Internet is to be committed to help customers all over the world in order to transform aspirations into achievements by harnessing the Internet's efficiency, speed and ubiquity [38].

3.4.3. Applications over Internet

The programs which can be written to do the job of a current manual task, or virtually anything, and which perform their job on the server (cloud server) via an internet connection rather than the traditional model of a program that has to be installed and run on a local computer are the Cloud Applications, or as a scientific definition Applications over Internet. Some examples of powerful programs which run in the cloud and they perform incredible feats of computing for the oblivious user who only needs an internet connection and a browser, are google applications, internet banking, and Facebook [38].

3.4.4. Energy Efficiency

As a definition, the Energy Efficiency is a way of managing and restraining the growth in energy consumption. By delivering more services for the same energy input or for the same services for less energy input may be something more energy efficient. As an example, when a Compact Florescent Light (CFL) bulb uses less energy (1/3 to 1/5) than an incandescent bulb to produce the same amount of lights, the Compact Florescent Light (CFL) is considered to be more energy efficient [39].

3.4.5. Computationally Capable

The services of computational clouds are leveraging the computationally intensive and ubiquitous mobile applications which have been enabled by the technology of Mobile Cloud Computing. Thus, a system is considered as computationally capable when it meets the requirements to provide us the results we want, by making the right calculations [38].

3.5 Chapter Summary

Considering that the Big Data is a new technology which develops rapidly in the field of telecommunications, and especially the modern field of wireless telecommunications, we have tried to combine this technology with the technologies of Mobile Cloud Computing and Internet of Things. Additionally, the technologies of Mobile Cloud Computing and Internet of Things begin to develop rapidly within the technology of wireless networks. We combined the MCC and the IoT with the technology of the Big Data, so we can check the common features and discover the benefits of these two technologies regarding their use with the Big Data Applications.

In this paper, we present a survey of Internet of Things Technology, with an explanation of its operation and use. Moreover, we present the main features of the Mobile Cloud Computing and its trade offs. Also, we have a presentation of the Big Data Applications and some of its basic features. Finally, we present the contribution of the Internet of Things technology, and the Mobile Cloud technology to Big Data Applications.

The exploration of the contribution provided by Internet of Things features, and by Mobile Cloud Computing features in dealing with the basic characteristics of the Big Data Applications, is shown in Table 3.1 and Table 3.2, respectively. However, based on these two tables, we can observe that the Internet of Things mostly contributes to the field of Network Partitioning and Unified Ethernet Fabrics of Big Data Applications. The Mobile Cloud Computing technology mostly contributes to the field of Predictable & Efficient, Holistic Network and Network Partitioning of Big Data Applications.

In conclusion, we can infer from the information which can be observed in Table 3.1 and Table 3.2 that the technologies of Internet of Things and Mobile Cloud Computing possess features which could be beneficial for the use of Big Data Applications. As for future research, we suggest that the use of the Big Data Applications is combined with the technologies of Internet of Things and Mobile Cloud Computing in order to achieve better results. Also, as a continuation of this research we will further examine the features of Big Data Applications which could be improved from the contribution of the technologies of Mobile Cloud Computing and Internet of Things.

3.6 Chapter References

[10] P. Bahl, R. Y. Han, L. E. Li, M. Satyanarayanan, “Advancing the State of Mobile Cloud Computing”, in Proceedings of the third ACM workshop on Mobile cloud computing and services (MCS ‘12), pp. 21-28, June 2012, Low Wood Bay Lake District, UK. [DOI: 10.1145/2307849.2307856]

[11] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, I. Brandic, “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility”, Elsevier, Future Generation Computer Systems, vol. 25, Issue: 6, pp. 599-616, June 2009. [DOI: 10.1016/j.future.2008.12.001]

[12] B. Prajapat, Dr. M. Shrivastava, “Mobile Cloud Computing through J2ME application: Cloud Enabled Web Services”, International Journal of Advanced Computer Research, Vol. 2, No. 4, Issue: 6, December 2012.

[13] D. Sohini, D. Suddhasil, “Uncoupling in Services of Mobile Cloud Computing using Tuple Space Model: Design and Formal Specifications”, in Proceedings of the first international workshop on Mobile cloud computing & networking MobileCloud’13, 29 July 2013, Bangalore, India.

[15] S. Abolfazli, Z. Sanaei, E. Ahmed, A. Gani, R. Buyya, “Cloud-Based Augmentation for Mobile Devices: Motivation, Taxonomies, and Open Challenges”, IEEE Communications Surveys & Tutorials, vol. 16, issue: 1, pp. 337-368, First Quarter 2014.

[16] M. Shiraz, E. Ahmed, A. Gani, Q. Han, “Investigation on runtime partitioning of elastic mobile applications for mobile cloud computing”, Springer, The Journal of Supercomputing, vol. 67, pp. 84-103, January 2014. [DOI: 10.1007/s11227-013-0988-6]

[17] G. Ananthanarayanan, I. Menache, “Big data analytics systems”, Microsoft Research, Redmond WA, USA.

[18] N. Mishra, C.-C. Lin, H.-T. Chang, “A Cognitive Adopted Framework for IoT Big-Data Management and Knowledge Discovery Prospective”, SAGE Journals, International Journal of Distributed Sensor Networks, vol. 11, Issue: 10, October 2015. [DOI: 10.1155/2015/718390]

[19] G. Suciu, V. Suciu, A. Martian, R. Craciunescu, A. Vulpe, I. Marcu, S. Halunga, O. Fratu, “Big Data, Internet of Things and Cloud Convergence – An Architecture for Secure E-Health Applications”, Springer, Patient Facing Systems, 2015, vol. 39, Article No. 141, September 2015. [DOI: 10.1007/s10916-015-0327-y]

[20] O. Niggemann, J. R. Kinnebrew, H. Khorasgani, S. Volgmann, A. Bunte, G. Biswas, “Data-Driven Monitoring of Cyber-Physical Systems Leveraging on Big Data and the Internet-of-Things for Diagnosis and Control”, in Proceedings of 26th International Workshop on Principles of Diagnosis, pp. 185-192, 31 August – 3 September 2015, Paris, France.

[21] R. P. Minch, “Location Privacy in the Era of the Internet of Things and Big Data Analytics”, in Proceedings of 48th Hawaii International Conference on System Sciences, 2015, Boise State University.

[22] M. Strohbach, H. Ziekow, V. Gazis, N. Akiva, “Towards a Big Data Analytics Framework for IoT and Smart City Applications”, Book Chapter, Springer, Modeling and Processing for Next-Generation Big-Data Technologies, vol. 4, pp. 257-282, 2015. [DOI: 10.1007/978-3-319-09177-8_11]

[23] “Internet of Things Global Standards Initiative”, ITU, Twelfth and last event Geneva, 14-20 July 2015. [Online]. Available: https://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx. [Accessed 26/06/2015]

[24] J. Höller, V. Tsiatsis, C. Mulligan, S. Karnouskos, S. Avesand, D. Boyle, “From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence”, Elsevier, 2014. [ISBN 978-0-12-407684-6]

[26] J. M. Batalla, “Advanced Multimedia Service Provisioning based on efficient interoperability of adaptive streaming protocol and High Efficient Video Coding”, Springer, Journal of Real-Time Image Processing, vol. 12, pp. 443-454, March 2015. [DOI: 10.1007/s11554-015-0496-4]

[27] D. Huang, “Mobile cloud computing,” IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, vol. 6, no. 10, pp. 27–31, 2011.

[28] P. Viswanathan, “Mobile Devices Expert, Cloud Computing – Is it Really All That Beneficial?”, Advantages and Disadvantages of Cloud Computing, [Online], , Available: http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm [Accessed 24/01/2015]

[29] E. Almort, S. Andersson, “A study of the advantages & disadvantages of mobile cloud computing versus native environment”, Bachelor Thesis in Software Engineering, School of Computing, Blekinge Institute of Technology, Sweden, May 2013.

[30] Blog: Follow what’s happening at Get Cloud Services, Mobile Cloud Computing – Pros and Cons, [Online], Posted: 23/12/2014, Available:

https://www.getcloudservices.com/blog/mobile-cloud-computing-pros-and-cons [Accessed 24/01/2015]

[31] P. Murray, "Enterprise grade cloud computing," in Proceedings of Third Workshop on Dependable Distributed Data Management (WDDM '09), March 2009, Nuremberg Germany. [ISBN: 978-1-60558-462-1] [DOI: 10.1145/1518691.1518692]

[32] Council of July 2002: Directive 2002/58/ec concerning the processing of personal data and the protection of privacy in the electronic communications sector. European Parliament. Accessed November 20th, 2011. [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:NOT

[33] L. Youseff, M. Butrico, D. Da Silva, "Toward a unified ontology of cloud computing", in Proceedings of Grid Computing Environments Workshop 2008, GCE'08, pp. 1-10, 12-16 November 2008, Austin, TX, USA. [DOI: 10.1109/GCE.2008.4738443]

[34] E. Shi, Y. Niu, M. Jakobsson, and R. Chow, "Implicit authentication through learning user behavior", Springer, in Proceedings of International Conference on Information Security (ISC 2010), Information Security pp 99-11, 25-28 October 2010, Boca Raton, FL, USA. [DOI: 10.1007/978-3-642-18178-8_9]

[35] "Data, data everywhere", The Economist. [Online], Posted: 27/02/2010, Available: https://www.economist.com/special-report/2010/02/27/data-data-everywhere [Accessed 26/09/2015]

[36] M. Hilbert, P. López, "The World's Technological Capacity to Store, Communicate, and Compute Information", AAAS, Science, vol. 332, issue: 6025, pp. 60–65, April 2011. [DOI:10.1126/science.1200970]

[37] L. Borovick, R. L. Villars, "The Critical Role of the Network in Big Data Applications", IDC Analyze the Future, White Paper, Sponsored by: Cisco Systems pp. 1-12, April 2012.

[38] Md Whaiduzzaman, M. N. Haque, Md R. K. Chowdhury, A. Gani "A Study on Strategic Provision of Cloud Computing Services", The Scientific World Journal, pp. 1-8, June 2014.

[39] S. K. Garg, S. Versteeg, R. Buyya, "A framework for ranking of cloud computing services", Elsevier, Future Generation Computer Systems, vol. 29, Issue: 4, pp. 1012-1023, June 2013. [DOI: 10.1016/j.future.2012.06.006]

Chapter 4

Secure integration of IoT and Cloud Computing

C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, "Secure integration of IoT and Cloud Computing", Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI:10.1016/j.future.2016.11.031]

4.1 Introduction

Mobile Cloud Computing is a new technology that refers to an infrastructure where both data storage and data processing operate outside of the mobile device. Another recent technology is the Internet of Things. Internet of Things is a new technology that is growing rapidly in the field of telecommunications. More specifically, IoT is related to wireless telecommunications. The main goal of the interaction and cooperation between things and objects sent through the wireless networks is to fulfill the objective set to them as a combined entity. In addition, there is a rapid development of both technologies, Cloud Computing and the Internet of Things, regard the field of wireless communications. In this paper, I present a survey of IoT and Cloud Computing with a focus on the security issues of both technologies. Specifically, I combine the two aforementioned technologies (i.e Cloud Computing and IoT) to examine the common features, and in order to discover the benefits of their integration. Concluding, we present the contribution of Cloud Computing to IoT technology. Thus, it shows how Cloud Computing technology improves the function of the IoT. Finally, I survey the security challenges of the integration of IoT and Cloud Computing.

4.2 Related Work

For the purpose of this paper we study and analyze previous literature which has been published in the field of cloud computing and Internet of Things, and their integration. The following paragraphs present the papers which contributed significantly in our study.

To begin with, a survey of the different security risks that pose a threat to the cloud is presented in [10]. Also, in [10] was given a survey more specific to the different security issues that has emanated due to the nature of the service delivery models of a cloud computing system. Moreover, an exploration of the roadblocks and solutions to provide a trustworthy cloud computing environment presented in [11]. Cloud computing is an evolving paradigm with tremendous momentum, but its unique aspects exacerbate security and privacy challenges.

Concerning the integration of Internet of Things and Cloud Computing, there have been made some previous studies. A propose of a new platform for using cloud computing capacities for provision and support of ubiquitous connectivity and real-time applications and services for smart cities' needs is given in [12]. Additionally, a presentation of a framework for data procured from highly distributed, heterogeneous, decentralized, real and virtual devices (sensors, actuators, smart devices) that can be automatically managed, analyzed and controlled by distributed cloud-based services shown in [12]. In order to realize the full sharing, free circulation, on-demand use, and optimal allocation of various manufacturing resources and capabilities, the applications of the technologies of IoT and CC in manufacturing are investigated in [13]. Furthermore, a CC- and IoT-based cloud manufacturing (CMfg) service system (i.e. CCIoT-CMfg) and its architecture are proposed, and the relationship among CMfg, IoT, and CC is analyzed. And finally, the advantages, challenges, and future works for the application and implementation of CCIoT-CMfg are discussed in [13]. The [14] mainly focuses on a common approach to integrate the Internet of Things (IoT) and Cloud Computing under the name of CloudThings architecture. Also, in [14] review the state of the art for integrating Cloud Computing and the Internet of Things, and examine an IoT-enabled smart home scenario to analyze the IoT application requirements. At the end, the CloudThings architecture, a Cloud-based Internet of Things platform which accommodates CloudThings IaaS, PaaS, and SaaS for accelerating IoT application, development, and management proposed in [14]. Furthermore, a presentation and discussion about some of the integration challenges of Iot and Cloud Computing that must be addressed to enable an intelligent transportation system to address issues facing the transportation sector such as high fuel prices, high levels of CO_2 emissions, increasing traffic congestion, and improved road safety are shown in [15].

A presentation of an approach to the development of Smart Home applications by integrating Internet of Things (IoT) with Web services and Cloud computing are shown in [16]. The approach focuses on: (1) embedding intelligence into sensors and actuators using Arduino platform; (2) networking smart things using Zigbee technology; (3) facilitating interactions with smart things using Cloud services; (4) improving data exchange efficiency using JSON data format. Also, it is shown an implementation of three use cases to demonstrate the approach's feasibility and efficiency, i.e., measuring home conditions, monitoring home appliances, and controlling home access. The [17] presents a Cloud centric vision for worldwide implementation of Internet of Things. The key enabling technologies and application domains that are likely to drive IoT research in the near future are discussed. A Cloud implementation using Aneka, which is based on interaction of private and public Clouds is also presented in [17]. Finally, it concludes the IoT vision by expanding on the need for convergence of WSN, the Internet and distributed computing directed at technological research community. Internet of Things (IoT) becoming so pervasive that it is becoming important to integrate it with cloud computing because of the amount of data IoT's could generate and their requirement to have the privilege of virtual resources utilization and storage capacity, but also, to make it possible to create more usefulness from the data generated by IoT's and develop smart applications for the users. This type of integration is referred to as Cloud of Things in [18]. With IoTs, anything can become part of the Internet and generate data. Moreover, data generated needs to be managed according to its requirements, in order to create more valuable services. For the previous purpose, integration of IoTs with cloud computing is becoming very important. This new paradigm is termed as Cloud of Things (CoTs) and it is presented in [19]. The [20] focuses in the attention of the authors on the integration of Cloud and IoT, which is what we call the CloudIoT paradigm. Also, many works in literature have surveyed Cloud and IoT separately and, more precisely, their main properties, features, underlying technologies, and open issues in [20]. However, these works lack a detailed analysis of the new CloudIoT paradigm, which involves completely new applications, challenges, and research issues. The [21] focuses on some of the key challenges involved in CoT and the proposal of smart gateway based communication. Cloud of Things, requires smart gateway to perform the rich tasks and preprocessing, which sensors and light IoTs are not capable of doing. Finally, the [22] presents a survey of integration components: Cloud platforms, Cloud infrastructures and IoT Middleware. In addition, some integration proposals and data analytics techniques are surveyed as well as different challenges and open research issues are pointed out.

Finally, we study integration algorithms and methods about the aforementioned technologies. In [23] the authors focus on Fuzzy C-Means based segmentation algorithms because of the segmentation accuracy they provide. Furthermore, the algorithms which have been studied need long execution times. Also, the authors of [23] accelerate the execution time of these algorithms using Graphics Process Unit (GPU) capabilities. At the end, the authors reach the achievement performance enhancement by up to 8.9x without compromising the segmentation accuracy. The main aim of the [24] is to perform a review of the basic methods used for such techniques and finding the emerging trends of the research in this area. The authors of [24] primary focus on summarize some well-known methods of face recognition in video sequences for application in biometric security and enumerate the emerging trends. The [25] in order to address the challenge of the lack of investigating on effective and efficient evaluations and measurements for security and trustworthiness of various social media tools, platforms and applications, surveys the state-of-the-art of social media networks security and trustworthiness particularly for the increasingly growing sophistication and variety of attacks as well as related intelligence applications. Also, the authors of [25] highlighted a new direction on evaluating and measuring the fundamental and underlying platforms. Furthermore, the authors propose a hierarchical architecture for crowd evaluations based on signaling theory and crowd computing, which is essential for social media ecosystem.

Table 4.1 lists the findings and the concepts examined in each paper. In more detail, in Table 4.1 could observe independently for each related review that have been studied useful information related to the year which published, the exact authors, and as a conclude for each paper the problems and the solutions which they deal with.

Year	**Author**	**Problems**	**Solutions**
2010	H. Takabi et al **[11]**	• Unique aspects exacerbate security and privacy challenges of Cloud Computing.	• Explores the roadblocks and solutions to providing a trustworthy Cloud Computing environment.
2011	S. Subashini & V. Kavitha **[10]**	• How safe is a Cloud Computing environment is. • Enterprise customers are still reluctant to deploy their business in the cloud. • Security is one of the major issues which reduces the growth of cloud computing and	• A new model targeting at improving features of an existing model must not risk or threaten other important features of the current model. • Cloud service users need to be vigilant in understanding the risks of data breaches in

		complications with data privacy and data protection continue to plague the market.	this new environment. • Different security issues that has emanated due to the nature of the service delivery models of a cloud computing system.
2013	J. Gubbi et al **[17]**	• Fueled by the recent adaptation of a variety of enabling wireless technologies, the IoT has stepped out of its infancy and is the next revolutionary technology in transforming the Internet into a fully integrated Future Internet. • The need for data-on-demand using sophisticated intuitive queries increases significantly.	• A Cloud centric vision for worldwide implementation of Internet of Things. • Cloud implementation using Aneka, which is based on interaction of private and public Clouds. • Expanding on the need for convergence of WSN, the Internet and distributed computing directed at technological research community.
2013	G. Suciu et al **[12]**	• Cloud Computing and Internet of Things (IoT) are two of the most popular ICT paradigms. • The convergence between cloud computing and IoT has become a hot topic over the last few years.	• A new platform for using cloud computing capacities for provision and support of ubiquitous connectivity and real-time applications and services for smart cities' needs. • A framework for data procured from highly distributed, heterogeneous, decentralized, real and virtual devices that can be automatically managed, analyzed and controlled by distributed cloud-based services.
2013	J. Zhou et al **[14]**	• User with a novel means of communicating with the Web world through ubiquitous object-enabled networks presented by Internet of Things. • Cloud Computing enables a convenient, on demand and scalable network access to a shared pool of configurable computing resources..	• A common approach to integrate the Internet of Things (IoT) and Cloud Computing under the name of CloudThings architecture. • An IoT-enabled smart home scenario to analyze the IoT application requirements.
2013	M. Soliman et al **[16]**	• Smart Home minimizes user's intervention in monitoring home settings and controlling home appliances.	• An approach to the development of Smart Home applications by integrating Internet of Things (IoT) with Web services and Cloud computing..
2014	M. Aazam et al **[18]**	• Everything is going to be connected to the Internet and its data will be used for various progressive purposes. • Internet of Things (IoT) becoming so pervasive that it is becoming important to integrate it with cloud computing.	• IoT's and cloud computing integration is not that simple and bears some key issues. Those key issues along with their respective potential solutions have been highlighted.
2014	M. Aazam et al **[21]**	• Integration of Internet of Things with Cloud Computing is gaining importance, with the way the trend is going on in ubiquitous computing world. • Internet of Things (IoT) becoming so pervasive that it is becoming important to integrate it with cloud computing.	• Integration of IoT with Cloud Computing, referred here as Cloud of Things, requires smart gateway to perform the rich tasks and preprocessing, which sensors and light IoTs are not capable of doing. • Focuses on some of the key challenges involved in CoT and the proposal of smart gateway based communication.
2014	F. Tao et al **[13]**	• Internet of Things (IoT) and cloud computing (CC) have been widely studied and applied in many fields, as they can provide a new method for intelligent perception and connection from M2M, and on-demand use and efficient sharing of resources, respectively.	• A CC- and IoT-based cloud manufacturing (CMfg) service system and its architecture are proposed. • The advantages, challenges, and future works for the application and implementation of CCIoT-CMfg are discussed.

2015	A. Botta et al **[20]**	• Cloud computing and Internet of Things (IoT) are two very different technologies that are both already part of our life. • A novel paradigm where Cloud and IoT are merged together is foreseen as disruptive and as an enabler of a large number of application scenarios.	• Integration of Cloud and IoT, which is called the CloudIoT paradigm. • A new CloudIoT paradigm, which involves completely new applications, challenges, and research issues.
2015	J. A. Guerrero Ibanez et al **[15]**	• Performance of transportation systems is of crucial importance for individual mobility, commerce, and for the economic growth of all nations. • It is imperative to improve the safety and efficiency of transportation.	• Integration challenges of IoT and CC that must be addressed to enable an intelligent transportation system to address issues facing the transportation sector.
2016	M. Diaz et al **[22]**	• Internet of Things comprises many interconnected technologies like RFID and WSAN in order to exchange information. • The limitations of associated devices in the IoT require a technology like Cloud Computing to supplement this field.	• A survey of integration components: Cloud platforms, Cloud infrastructures and IoT Middleware.
2016	M. Aazam et al **[19]**	• It is becoming very difficult to manage power constrained small sensors and other data generating devices. • Data generated needs to be managed according to its requirements, in order to create more valuable services.	• Integration of IoTs with cloud computing is becoming very important – Cloud of Things. • CoTs provide means to handle increasing data and other resources of underlying IoTs and WSNs.
2016	M. Alsmirat et al **[23]**	• Big revolution in information technology that is used to diagnose many illnesses and saves patients lives. • Image segmentation is a mandatory step in many image processing based diagnosis procedures.	• Fuzzy C-Means based segmentation algorithms provide segmentation accuracy. • Accelerate the execution time of Fuzzy C-Means algorithms using Graphics Process Unit (GPU) capabilities.
2016	B. B. Gupta et al **[24]**	• Face recognition from video has gained attention due to its popularity and ease of use with security systems based on vision and surveillance systems. • Automated video based face recognition system provides a huge assortment of challenges as it is necessary to perform facial verification under different viewing conditions.	• Perform a review of the basic methods used for such techniques and finding the emerging trends of the research in this area. • Summarize some well-known methods of face recognition in video sequences for application in biometric security and enumerate the emerging trends.
2016	Z. Zhang et al **[25]**	• Social media security and trustworthiness issues have become increasingly serious. • Lack of investigating on effective and efficient evaluations and measurements for security and trustworthiness of various social media tools, platforms and applications.	• Survey on the state-of-the-art of social media networks security and trustworthiness particularly for the increasingly growing sophistication and variety of attacks. • Highlight a new direction on evaluating and measuring those fundamental and underlying platforms. • Propose a hierarchical architecture for crowd evaluations based on signaling theory and crowd computing.

Table 4.1: Mapping problems against referenced solutions.

4.3 Internet of Things

The Internet of Things is a network of devices that transmit, share, and use data from the physical environment to provide services to individuals, corporations, and society. The objects-things function either individually or in connection with other objects or individuals, and have unique IDs (identifiers). Also, the Internet of Things has different applications in health, transport, environment, energy or types of devices: sensors, devices worn/carried (wearable), e.g. watch, glasses, home automation (domotics).

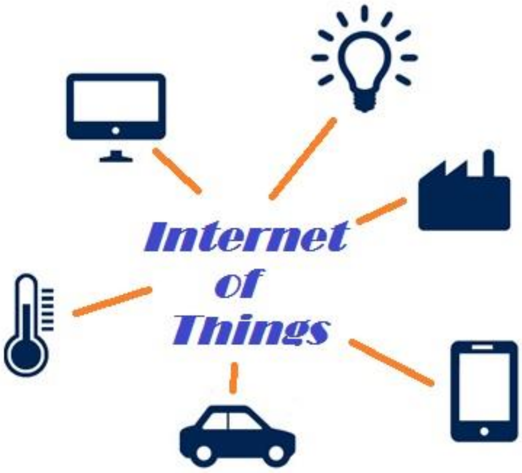


Figure 4.1: Internet of Things Technology.

4.3.1. Internet of Things: Advantages of the data

What does it mean when the devices and sensors are networked together and communicate with each other? How can the Internet of Things affect our daily life? GPS systems, alarm systems, and thermostats, all send and receive constant feeds to monitor and automate activities in our daily lives [26]. And the not so obvious: Mosaic, cups, clothes and other everyday objects can also join network to send and receive data over the Internet.

Opportunities where the streaming data will create new markets in order to inspire positive change or to enhance existing services are examined by businesses. Some examples of sectors that are at the heart of these developments are listed below [27]:

a) Smart solution in the bucket of transport: Smart solutions in the bucket of transport, achieve a reduction of traffic on the roads, reduce fuel consumption, set priorities in vehicle repair programs, and save lives.

b) Smart power grids incorporating more renewable: Smart power grids incorporating more renewables improve system reliability, and reduce the charges consumers, thus providing cheaper electricity.

c) Remote monitoring of patients: Remote monitoring of patients provides easy access to health care, improves the quality of services, increases the number of people served, and saves money.

d) Sensors in homes and airports: Sensors in homes and airports, or even in your shoes or doors, improve safety by sending signals when left unused for a certain period of time or when used in the wrong time.

e) Engine monitoring sensors that detect & predict maintenance issues: Engine monitoring sensors that detect and predict maintenance issues, improve inventory replenishment, and even define priorities in scheduling maintenance work, repairs, and regional operations.

4.3.2. Internet of Things Security

IoT security is the area of endeavor concerned with safeguarding connected devices and networks in the Internet of things. The Internet of Things involves the increasing prevalence of objects and entities – known, in this context as things -- provided with unique identifiers and the ability to automatically transfer data over a network. Much of the increase in IoT communication comes from computing devices and embedded sensor systems used in industrial machine-to-machine (M2M) communication, smart energy grids, home and building automation, vehicle to vehicle communication and wearable computing devices [28] [29].

The main problem is that because the idea of networking appliances and other objects is relatively new, security has not always been considered in product design. IoT products are often sold with old and unpatched embedded operating systems and software. Furthermore, purchasers often fail to change the default passwords on smart devices -- or if they do change them, fail to select sufficiently strong passwords. To improve security, an IoT device that needs to be directly accessible over the Internet, should be segmented into its own network and have network access restricted. The network segment should then be monitored to identify potential anomalous traffic, and action should be taken if there is a problem.

Security experts have warned of the potential risk of large numbers of unsecured devices connecting to the Internet since the IoT concept was first proposed in the late 1990s. In December of 2013, a researcher at Proofpoint, an enterprise security firm, discovered the first IoT botnet. According to Proofpoint, more than 25 percent of the botnet was made up of devices other than computers, including smart TVs, baby monitors and other household appliances [28].

4.3.3. Internet of Things Security model

In the field of Internet of Things technology there are System models and initial conditions considered are as similar as that of [30]. A wireless network model with a source-destination pair, N trusted relays and J eavesdroppers $(J \leq 1)$ are considered. Assume that the global CSE is available. The eavesdropper channel, source encoding schemes, decoding schemes and cooperative protocol are considered to be public, only source message is assumed to be confidential. In this paper, the discussion is limited to two main cooperative schemes: decode-and-forward (DF) and amplify-and-forward (AF) [31].

Notation	***Description***
$(.)*$	Conjugate
$(.)^T$	Transpose
$(.)^{\dagger}$	Conjugate transpose
I_M	Identity matrix of size M x N
$diag\{a\}$	Diagoanl matrix with the elements of vector α along its diagonal
$\lVert\alpha\rVert$	2-norm of vetor α
$0_{M\times N}$	All-zero matrix of size M x N
$\log(.)$	Base-2 logarithm
h_{SD}^{*}	Baseband complex channel gain between source and destination
h_{SE}^{*}	Channel vector(J x 1) between source and J eavesdroppers
h_{SR}^{*}	Channel vector(N x 1) between source and N relays
h_{RD}^{*}	Channel vector(N x 1) between N relays and destination
H_{RE}^{*}	Channel vector(N x J) between N relays and J eavesdroppers

Table 4.2: Notations used for the main cooperative schemes DF and AF.

Decode-and-forward (DF)

There are two main stages in DF. Source broadcasts its encoded symbols to its trusted relays using the first transmission slot in Stage 1. When transmitting the symbol *x*, the received signals at the *N* relays are given by,

$$y_r = \sqrt{P_s} h_{SR}^{*} x + n_r \quad (1)$$

where P_s is the transmit power of source and n_r is the noise vector at relays [31].

In Stage 2, all the trusted relays that successfully decode the message, re-encode the message and cooperatively transmit the re-encoded symbols to the destination by using the second transmission slot. Each relay transmits a weighted version of the re-encoded symbol. When transmitting the symbol $\tilde{x}$, the received signal at the destination is given by,

$$y_d = h_{RD}^{\dagger} w \tilde{x} + n_d \quad (2)$$

while the received signal at the eavesdroppers is expressed in vector form as,

$$y_e = H_{RE}^{\dagger} w \tilde{x} + n_e \quad (3)$$

The transmit power budget for Stage 2 is considered to be P - P_s where P is the total power for transmitting one symbol and P_s is the transmit power of source [31].

Amplify-and-forward (AF)

AF is also a two-stage scheme as that of DF. Stage 1 is the same for both AF and DF, except that the transmit power can be different. The trusted relays forward the signals that are received during Stage 1 to the destination, using the second transmission slot in Stage 2. That is, each relay transmits a weighted version of the noisy signal that they received during Stage 1. The transmitted signals of all relays are denoted by the product of $diag\{w\}y_r$, where w is the weight vector and y_r is given by (1). The received signal at the destination is given by [30],

$$y_d = \sqrt{P_s}\, h_{RD}^{\dagger} diag\{w\} h_{SR}^{*} x + h_{RD}^{\dagger} diag\{w\} n_r + n_d \quad (4)$$

The received signals at the eavesdroppers, in a vector form, is denoted by [26],

$$y_e = \sqrt{P_s}\, H_{RE}^{\dagger} diag\{w\} h_{SR}^{*} x + H_{RE}^{\dagger} diag\{w\} n_r + n_e \quad (5)$$

where P_s is the transmit power of source, n_r is the noise vector at relays and x is the received signal. Also, equations (4) and (5) generated from (1) and (2), and (1) and (3) respectively.

Additionally, another security challenge in IoT is the encryptions algorithm. The RSA algorithm, which is the most commonly used public key algorithm in the Internet, can be used in sensor networks with the assistance of a Trusted Platform Module (TPM), which costs less than 5% of a common sensor node [32]. Thus, the memory has been measured for a fully authenticated handshake with 2048-bit RSA keys. This type of handshake has the largest memory requirements since it needs more code and buffer space for the client's *Certificate* and *CertificateVerify* messages. The memory increased its use because the code basically contains hundreds of statements form *buffer[x] = 0xff*. The use of this encryption algorithm in IoT's security could provide better communication privacy in its functionality.

4.4 Cloud Computing

Cloud computing provides computing, storage, services, and applications over the Internet. In general, to render smartphones energy efficient and computationally capable, major changes to the hardware and software level are required. This entails the cooperation of developers and manufacturers. [33]. Mobile cloud computing is defined as an integration of cloud computing technology with mobile devices in order to make the mobile devices resource-full in terms of computational power, memory, storage, energy, and context awareness. The technology of Mobile Cloud computing is the outcome of interdisciplinary approaches combining mobile computing with cloud computing. Thus, this transdisciplinary domain is also referred as mobile cloud computing [33].

There are two perspectives in which the term Mobile Cloud refers: a) infrastructure based, and b) ad-hoc mobile cloud. In the infrastructure based mobile cloud, the hardware infrastructure remains static and also provides services to the mobile users. Nevertheless, there are several applications which utilize cloud resources, but the usage is limited to only storage and application-specific services such as Apple’s Siri (voice based personal assistant) and iCloud storage service.

Figure 4.2: Cloud Computing Technology.

4.4.1. Cloud Computing Features

As all technologies, so the Cloud Computing technology has some features which determine its function. These features are analyzed and outlined subsequently.

<u>Storage over Internet</u>

Storage over Internet can be defined as a technology framework that uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks to link servers and storage devices, and to facilitate storage solution deployment. The Storage over Internet technology is also known as Storage over Internet Protocol (SoIP) technology.

With the combination of the best storage and networking industry approaches, SoIP provides high-performance and scalable IP storage solutions [34] [35] [36].

Service over Internet

The main objective of the Service over Internet is to be committed to help customers all over the world in order to transform aspirations into achievements by harnessing the Internet's efficiency, speed and ubiquity [34] [35].

Applications over Internet

The programs which can be written to do the job of a current manual task, or virtually anything, and which perform their job on the server (cloud server) via an internet connection rather than the traditional model of a program that has to be installed and run on a local computer are the Cloud Applications, or as a scientific definition Applications over Internet. Some examples of powerful programs which run in the cloud and they perform incredible feats of computing for the oblivious user who only needs an internet connection and a browser, are google applications, internet banking, and Facebook [34] [35] [37].

Energy Efficiency

As a definition, the Energy Efficiency is a way of managing and restraining the growth in energy consumption. By delivering more services for the same energy input or for the same services for less energy input may be something more energy efficient. As an example, when a Compact Florescent Light (CFL) bulb uses less energy (1/3 to 1/5) than an incandescent bulb to produce the same amount of lights, the Compact Florescent Light (CFL) is considered to be more energy efficient [34] [35] [37].

Computationally Capable

The services of computational clouds are leveraging the computationally intensive and ubiquitous mobile applications which have been enabled by the technology of Mobile Cloud Computing. Thus, a system is considered as computationally capable when it meets the requirements to provide us the results we want, by making the right calculations [34] [35].

4.4.2. Mobile Cloud Computing trade offs

Mobile Cloud Computing has some disadvantages-limitations which should be eliminated over the years in order to achieve a better and more ideal use. A number of businesses and especially the smaller ones need to be aware of these limitations before going in for this technology.

Security

One major issue of the Mobile Cloud Computing is the security issue. Before someone adopts this technology, they should know that all the company's sensitive information would be surrender to a third-party cloud service provider. This could potentially put the company in great risk. Hence, someone must be absolutely sure that they would choose the most reliable service provider, who will keep the information completely safe [38] [39].

Connectivity

Internet connection is critical to Mobile Cloud Computing. Thus, the user should be certain that there is a good result before opting for these services. Since someone owes a mobile device which is connected to the internet has become the norm in the wireless world of today, Mobile Cloud Computing has a very large potential user base [40].

Performance

Another major concern of the Mobile Cloud Computing pertains to its performance. Some users feel performance is not as good as in native applications. Thus, checking with one service provider and understanding their track record is advisable [41] [42].

Latency (Delay)

In mobile cloud computing, latency (sometimes referred as turnaround time) is defined as the time involved in offloading the computation and getting back the results from the nearby infrastructure or cloud.

Privacy

Data privacy is important and is one of the main bottlenecks that restrict consumers from adopting mobile cloud computing. Therefore, to gain consumers trust in the mobile cloud, the application models must support application development with privacy protection, and implicit authentication mechanisms [39] [43].

4.4.3. Mobile Cloud Computing Security Issues

Cloud computing security or cloud security is an evolving sub-domain of computer security, network security, and, more broadly, information security. It refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing.

Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in third-party data centers [44]. Organizations use the Cloud in a variety of different service models (SaaS, PaaS, and IaaS) and deployment models (Private, Public, Hybrid, and Community) [45] [46]. There are a number of security concerns associated with cloud computing. These issues fall into two broad categories: security issues faced by cloud providers (organizations providing software, platform, or infrastructure-as-a-service via the cloud) and security issues faced by their customers (companies or organizations who host applications or store data on the cloud) [38] [48]. The responsibility is shared, however. The provider must ensure that their infrastructure is secure and that their clients' data and applications are protected while the user must take measures to fortify their application and use strong passwords and authentication measures.

4.4.4. Cloud Computing Security model

In order to provide secure communication over the network, encryption algorithm plays an important role. It is a valuable and fundamental tool for the protection of the data. Encryption algorithm converts the data into scrambled form by using “a key” and only the user have the key to decrypt the data. Regarding the researches that have been made, an important encryption technique is the Symmetric key Encryption. In Symmetric key encryption, only one key is used to encrypt and decrypt the data. In this encryption technique the most used algorithm is the AES [49] [50].

AES (Advanced Encryption Standard) is the new encryption standard recommended by NIST to replace DES algorithm. Brute force attack is the only effective attack known against it, in which the attacker tries to test all the characters combinations to unlock the encryption. The AES algorithm block ciphers. It has variable key length of 128, 192, or 256 bits; default 256. It encrypts data blocks of 128 bits in 10, 12 and 14 round depending on the key size. AES encryption is fast and flexible; it can be implemented on various platforms especially in small devices. Also, AES has been carefully tested for many security applications [51] [52].

A part of the AES algorithm represented in this work. This algorithm uses the original key consists of the number of bytes in any case, which are represented as a 4x4 matrix.

Algorithm - part of the AES algorithm

```
Cipher(byte[] input, byte[] output)
{
    byte[4,4] State;
    copy input[] into State[] AddRoundKey
    for (round = 1; round < Nr-1; ++round)
    {
        SubBytes ShiftRows MixColumns AddRoundKey
    }
    SubBytes ShiftRows AddRoundKey
    copy State[] to output[]
}
```

AES algorithm considered as better than others for a number of reasons, which is follows [53]:

- ✓ AES performs consistently well in both hardware and software platforms under a wide range of environments. These include 8-bit and 64-bit platforms and DSP’s.
- ✓ Its inherent parallelism facilitates efficient use of processor resources resulting in very good software performance.
- ✓ This algorithm has speedy key setup time and good key agility.

- ✓ It requires less memory for implementation, making it suitable for restricted-space environments.
- ✓ The structure has good potential for benefiting from instruction-level parallelism.
- ✓ There are no serious weak keys in AES.
- ✓ It supports any block sizes and key sizes that are multiples of 32 (greater than 128-bits).
- ✓ Statistical analysis of the cipher text has not been possible even after using huge number of test cases.
- ✓ No differential and linear cryptanalysis attacks have been yet proved on AES.

Additionally, there is an important encryption technique from the Asymmetric key Encryption. In Asymmetric key encryption, two keys, private and public keys, are used. Public key is used for encryption and private key is used for decryption [49] [50].

RSA is an Internet encryption and authentication system that uses an algorithm developed in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman. The RSA algorithm is the most commonly used encryption. Till now it is the only algorithm used for private and public key generation and encryption. It is a fast encryption [54].

The RSA algorithm which studied in this work uses a key generator that provides two large primes. Those primes are used in order to procced the encryption mode. The two large primes represent the two types of keys that we use in decryption and encryption, the public key and the secret key.

Algorithm - RSA algorithm

```
Key Generation: KeyGen(p, q)
Input: Two large primes – p, q
Compute $n = p . q$
            $\varphi (n) = (p - 1)(q - 1)$
    Choose e such that $gcd(e, \varphi (n)) = 1$

Determine d such that $e . d \equiv 1 \mod \varphi (n)$
Key:
    public key = (e, n)
    secret key= (d, n)
Encryption:
    c = me mod n
where c is the cipher text and m is the plain text.
```

RSA has a multiplicative homomorphic property i.e., it is possible to find the product of the plain text by multiplying the cipher texts. The result of the operation will be the cipher text of the product [44].

The equation given $c_i = E(mi) = mi^e \; mod \; n$, then we have the following:

$$(c1 \,.\, c2) \bmod n = (m1 \,.\, m2)^e \bmod n$$

4.5 IoT and Cloud Computing Integration

Moreover, a new generation of services, based on the concept of the 'cloud computing', has made its appearance in the last few years with the purpose of providing access to the information and the data from any place at any time, thus restricting or eliminating the need for hardware equipment. The term 'cloud computation' is defined as the use of computing logistical resources, as well as the software level, through the use of services transported over the Internet. Nowadays, cloud computing services comprise one of the world's largest areas of competition between giant companies in the IT sector and software [55]. Cloud Computing is a technology which can be set as a base technology in the use of IoT.

More specifically, Mobile Cloud Computing is defined as an integration of cloud computing technology with mobile devices so as to make the mobile devices resourceful in terms of computational power, memory, storage, energy, and context awareness. Mobile Cloud Computing is the outcome of interdisciplinary approaches, combining mobile computing and cloud computing [56]. In addition, Cloud computing provides computing, storage, services, and applications over the Internet. The technology of Mobile Cloud Computing is the outcome of interdisciplinary approaches, combining mobile computing with cloud computing. Thus, this transdisciplinary domain is also referred as Mobile Cloud Computing [33].

Some of the main features of the Cloud Computing technology which relate to the characteristics of both Internet of Things are: a) Storage over Internet, b) Service over Internet, c) Applications over internet, d) Energy efficiency and e) Computationally capable. Tables 4.2 lists the features of Mobile Cloud Computing regarding the convenience this technology offers when combined with the characteristics of IoT.

Internet of Things characteristics	Storage over Internet	Service over Internet	Applications over Internet	Energy efficiency	Computationally capable
Smart solution in the bucket of transport	X	X	X		X
Smart power grids incorporating more renewable	X	X		X	X
Remote monitoring of patients		X	X		X
Sensors in homes and airports	X	X	X	X	X
Engine monitoring sensors that detect & predict maintenance issues		X	X	X	X

Table 4.3: Contributions of Cloud Computing in Internet of Things.

Table 4.3 lists the features of Cloud Computing technology regarding the convenience this technology offers. Also, it enumerates the main features of the Internet of Things technology. The main purpose of Table 4.2 is to show which of the specific features of Cloud Computing technology, related more and improve the features of Internet of Things technology. As we can observe from Table 4.2, the feature of IoT which affected more by the features of Cloud Computing is “Sensors in homes and airports”. Regarding the Cloud Computing, the feature which affected more are “Service over Internet” and “Computationally capable”. As a general conclusion, we can observe that those two technologies contribute more each other in many of their features.

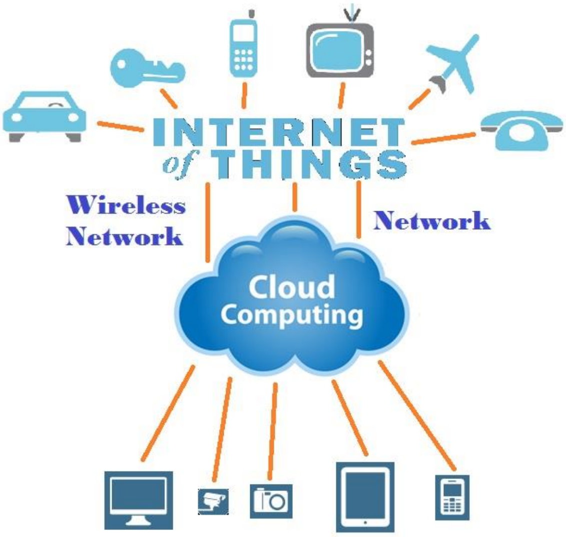


Figure 4.3: IoT & Cloud Computing Integration.

Through the integration of IoT and Cloud we have the opportunity to expand the use of the available technology that provided in cloud environments. Applications and information that use the Internet of Things technology with this integration can be used through the cloud storage. The integration of IoT and Cloud technologies represented in Figure 4.3. The cloud offers to mobile and wireless users to access all the information and the application that needed for the IoT connectivity.

4.5.1 Security issues in IoT and Cloud Computing integration

There is a rapid and independent evolution considering the two words of IoT and Cloud Computing. To begin with, the virtually unlimited capabilities and resources of Cloud Computing in order to compensate its technological constrains, such as processing, storage and communication, could be a benefit for the Internet of Things technology [58]. Also, the IoT technology extends its scope to deal with real world things in a more distributed and dynamic manner and by delivering new services in a large number of real life scenarios, might be beneficial for the use of Cloud Computing technology. In many cases, Cloud can provide the intermediate layer between the things and the applications, hiding all the complexity and functionalities necessary to implement the latter [20].

Through the integration of IoT and Cloud Computing could be observed that Cloud Computing can fill some gaps of IoT such the limited storage and applications over internet. Also, IoT can fill some gaps of Cloud Computing such the main issue of limited scope. Based in motivations such those referred previously and the important issue of security in both technologies we can consider some drivers for the integration. The security issue of this integration has a serious problem. When critical IoT applications move towards the Cloud Computing technology, concerns arise due to the lack of trust in the service provider or the knowledge about service level agreements (SLAs) and knowledge about the physical location of data. Consequently, new challenges require specific attention as mentioned in surveys [59] [60] [61]. Multi-tenancy could also compromise security and lead to sensitive information leakage. Moreover, public key cryptography cannot be applied at all layers due to the computing power constraints imposed by the things. These are examples of topics that are currently under investigation in order to tackle the big challenge of security and privacy in Cloud Computing and IoT integration [20].

Subsequently, some challenges about the security issue in the integration of two technologies are listed [20].

a) Heterogeneity. A big challenge in Cloud Computing and IoT integration is related to the wide heterogeneity of devices, operating systems, platforms, and services available and possibly used for new or improved applications [62].

b) Performance. Often Cloud Computing and IoT integration's applications introduce specific performance and QoS requirements at several levels (i.e. for communication, computation, and storage aspects) and in some particular scenarios meeting requirements may not be easily achievable [63] [64].

c) Reliability. When Cloud Computing and IoT integration is adopted for mission-critical applications, reliability concerns typically arise e.g., in the context of smart mobility, vehicles are often on the move and the vehicular networking and communication is often intermittent or unreliable. often intermittent or unreliable. When applications are deployed in resource constrained environments a number of challenges related to device failure or not always reachable devices exists [65].

d) Big Data. With an estimated number of 50 billion devices that will be networked by 2020, specific attention must be paid to transportation, storage, access, and processing of the huge amount of data they will produce. The ubiquity of mobile devices and sensor pervasiveness, indeed call for scalable computing platforms [66].

e) Monitoring. As largely documented in the literature, monitoring is an essential activity in Cloud environments for capacity planning, for managing resources, SLAs, performance and security, and for troubleshooting [67].

IoT & Cloud Computing security challenges	Heterogeneity	Performance	Reliability	Big Data	Monitoring
Internet of Things		X	X	X	X
Cloud Computing	X	X		X	

Table 4.4: Affects of IoT & Cloud Computing security challenges.

Table 4.4 lists the two technologies that we study in this paper and the challenges of their integration that arising from our study. These challenges related to the security issue in the integration of two aforementioned technologies and they listed in detailed in subsection 4.5.1. As we can observe from Table 4.4, the both technologies have two common main challenges of their integration which are Performance and Big Data. Additionally, we can observe that Internet of Things technology related to more challenges (4) than the Cloud Computing technology (3).

4.5.2. Proposed Efficient IoT and Cloud Computing security model

As we can infer, by taking advantage of the reasons which AES algorithm provides better secure in Cloud Computing and the two models that give benefits in security issues in IoT we can propose a new method that uses those benefits in order to improve the security and privacy issues in the integration of two technologies.

The AES algorithm provides the ability to have speed key setup time a good key agility. So, if we use this algorithm in the functionality of DF model, we could have a trusted relay method with an encryption of a speed key setup. Therefore, instead the trust relay use that DF and AF methods provide we can seize also there no serious weak keys in AES and so we could have a beneficial security use of the encryption in the integrated new model. Moreover, we can take advantage the less memory which AES needs for implementation that makes it for restricted-space environments. Thus, we can seize the transmit power that the AF model provides and as a result we can have a better and more trusted transmission. In the way of transmission, when the symbol transmitted with the use of DF model, the received signal at destination is given by the equation (2), which mentioned in previous section.

With this proposed model we can extend the advances of Internet of Things and Cloud Computing, by developing a highly innovative and scalable service platform to enable secure and privacy services. Through this research we can propose the following algorithm which extends the security advances of both technologies.

```
Key Generation: KeyGen(p, q)
Input: Two large primes – p, q
Compute $n = p . q$
        $buffer(n) = (p - 1)(q - 1)$
  Choose e such that $gcd(e, buffer(n)) = 1$
```

In which algorithm the equation method that contains hundreds of statements of the form *buffer[x]=0xff* is combined. With the use of this new type of RSA algorithm in the encryption process, we can conclude that a higher level of communications' security can be provided in the functionalities of the IoT.

Key:
public key = (e, n)
secret key= (d, n)
Encryption:
c = me mod n

where *c* is the cipher text and m is the plain text.

Also, as a proposal of this work could be the following part of algorithm which uses the original key consists of 128 bits/16 bytes which are represented as a 4x4 matrix. With the use of this part of AES algorithm we can draw that data which encrypted with 128bit (or 16 bytes) can be have better encrypted as an 4x4 matrix in order of providing a better use of communication privacy.

```
Cipher(byte[] input, byte[] output)
{
        byte[4,4] State;
        copy input[] into State[] AddRoundKey
        for (i=4; i<44; i++)
        {
                T = W[i-1];
                if (i mod 4 = = 0)
                        T = Substitute (Rotate (T)) XOR RConstant [i/4];
                W[i] = W[i-4] XOR T;
                SubBytes ShiftRows MixColumns AddRoundKey
        }
        SubBytes ShiftRows AddRoundKey
        copy State[] to output[]
}
```

Characteristics	Developed	Key length	Rounds	Certifications	Speed
AES	1998	128, 192 or 256 bits	10, 12 or 14	AES winner, CRYPTREC, NESSIE, NSA	Very fast
RSA	1977	1024-4096 bits	1	PKCS#1, ANSI X9.31, IEEE 1363	Very fast

Table 4.5: Comparison of AES and RSA algorithms.

Table 4.5 lists the key characteristics of the two encryption algorithms which have been studied and used in order to use them for the experimental proposal. The key characteristic which is more important is their Speed in which both algorithms are very fast. The key characteristic in which there is a relative difference is the Rounds, where AES needs 10, 12 or 14 rounds instead of the RSA that needs only 1.

4.3. Experimental results

Considering the benefits of the security models and algorithms of Internet of Things and Cloud Computing technologies we can observe that we can have a beneficial use of integration those two technologies. Instead of the wide use of IoT we can take advantage that Cloud Computing security through the AES algorithm performs consistently well in both hardware and software platforms under a wide range of environments. This use could be possible for all type of platforms and DSPs. Furthermore, the new integrated technology could has good potential for benefiting from instruction-level parallelism and will support any type of block sizes and key sizes that are multiples of 32 and used both of IoT and Cloud Computing. Also, each transmitted signal through the new technology can transmitted as a relay and trusted signal with a weighted version of the re-encoded symbol. By the use of RSA algorithm we can take advantage the two keys encryption in order to provide better secure in the use of the new model.

Through this integration we can achieve some useful functions, i.e. we can use the Cloud-based IoT service in order to connect sensors and also made them capable to share the sensor readings with others, reducing the security issues. Furthermore, another useful function is that we can use the HTTP protocol in order to send data between IoT things and the Cloud Computing applications. Moreover, some of the key advantages and challenges that can be defined from this integration are: 1) Both the physical hardware manufacturing resource and software manufacturing can be intelligently perceived and connected into the wider networks with the support of IoT technologies. 2) The collected information and data can be communicated and transmitted between M2M under the support of specific IoT technologies. 3) The collected and transmitted information can be processed and computed according to specific requirements under the support of different Cloud Computing service, and some useful data and decision information can be intelligently generated and obtained.

However, many other challenges and other benefits remains to be addressed through the integration of Internet of Things and Cloud Computing regarding the security issues, but also regarding the hole use of both technologies together.

AES Characteristics	Key length	Rounds	Certifications	Speed
Internet of Things	X		X	X
Cloud Computing	X	X	X	
IoT & CC integration	X	X	X	X

Table 4.6: AES contribution in IoT and Cloud Computing.

RSA Characteristics	Key length	Rounds	Certifications	Speed
Internet of Things	X		X	X
Cloud Computing	X	X	X	
IoT & CC integration	X	X	X	X

Table 4.7: RSA contribution in IoT and Cloud Computing

The Tables 4.6 and 4.7 exhibiting the key characteristic of the two encryption algorithms that used in order to achieve integration of the technologies of IoT and Cloud Computing concerning the security issue. Table 4.6 presents which of the key characteristics of AES encryption algorithm contributes both IoT and Cloud Computing technologies, and at the end how completely contributes the integration model of IoT and Cloud Computing. Subsequently, Table 4.7 presents which of the key characteristics of RSA encryption algorithm also contributes both IoT and Cloud Computing technologies, and at the end how completely contributes the integration model of IoT and Cloud Computing too.

4.6 Chapter Summary

The Cloud Computing technology offers many possibilities, but also places several limitations as well. Cloud Computing refers to an infrastructure where both the data storage and the data processing happen outside of the mobile device. In this paper, we present a survey of Internet of Things Technology, with an explanation of its operation and use. Moreover, we present the main features of the Cloud Computing and its trade offs. Cloud Computing refers to an infrastructure where both data storage and data processing happen outside of the mobile device. Also, the Internet of Things is a new technology which is growing rapidly in the field of telecommunications, and especially in the modern field of wireless telecommunications.

The main goal of the interaction and cooperation between things and objects sent through the wireless networks is to fulfil the objective set to them as a combined entity. In addition, based on the technology of wireless networks, both the technologies of Cloud Computing and Internet of Things develop rapidly. In this paper, we present a survey of IoT and Cloud Computing with a focus on the security issues of both technologies. Specifically, we combine the two aforementioned technologies (i.e Cloud Computing and IoT) in order to examine the common features, and in order to discover the benefits of their integration. Concluding, the contribution of Cloud Computing to the technology IoT, and it shows how the Cloud Computing technology improves the function of the IoT was presented. At the end, the security challenges of the integration of IoT and Cloud Computing were surveyed through the proposed algorithm model, and also there is a presentation of how the two encryption algorithms which were used contributes in the integration of IoT and Cloud Computing. This can be the field of future research on the integration of those two technologies. Regarding the rapid development of both technologies the security issue must be solved or reduced to a minimum in order to have a better integration model. These security challenges that surveyed in this paper could be the sector for further research as a case study, with the goal of minimizing them.

4.7 Chapter References

[10] S. Subashini, V. Kavitha, “A survey on security issues in service delivery models of cloud computing”, Elsevier, Journal of Network and Computer Applications, vol. 34, Issue: 1, pp. 1-11, January 2011. [DOI: 10.1016/j.jnca.2010.07.006]

[11] H. Takabi, J. B. D. Joshi, “Security and Privacy Challenges in Cloud Computing Environments”, IEEE Security & Privacy, vol. 8, Issue: 6, pp. 24-31, November-December 2010. [DOI: 0.1109/MSP.2010.186]

[12] G. Suciu, A. Vulpe, S. Halunga, O. Fratu, G. Todoran, V. Suciu, “Smart Cities Built on Resilient Cloud Computing and Secure Internet of Things”, in Proceedings of 19th International Conference on Control Systems and Computer Science 2013, 29-31 May 2013, Bucharest, Romania. [DOI: 10.1109/CSCS.2013.58]

[13] F. Tao, Y. Cheng, L. Zhang, B. H. Li, “CCIoT-CMfg: Cloud Computing and Internet of Things-Based Cloud Manufacturing Service System”, IEEE Transactions on Industrial Informatics, vol. 10, Issue: 2, pp. 1435-1442, May 2014. [DOI: 10.1109/TII.2014.2306383]

[14] J. Zhou, T. Leppanen, E. Harjula, M. Ylianttila, T. Ojala, C. Yu, H. Jin, L. T. Yang, “CloudThings: a Common Architecture for Integrating the Internet of Things with Cloud Computing”, in Proceedings of the 2013 IEEE 17th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 27-29 June 2013, Whistler, BC, Canada. [DOI: 10.1109/CSCWD.2013.6581037]

[15] J. A. Guerrero-Ibáñez, S. Zeadally, J. Contreras-Castillo, “Integration Challenges of Intelligent Transportation Systems with Connected Vehicle, Cloud Computing, and Internet of Things Technologies”, IEEE Wireless Communications, vol. 22, Issue: 6, pp. 122-128, December 2015. [DOI: 10.1109/MWC.2015.7368833]

[16] M. Soliman, T. Abiodun, T. Hamouda, J. Zhou, C.-H. Lung, “Smart Home: Integrating Internet of Things with Web Services and Cloud Computing”, in Proceedings of IEEE 5th International Conference on Cloud Computing Technology and Science, 2-5 December 2013, Bristol, UK. [DOI: 10.1109/CloudCom.2013.155]

[17] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions”, Elsevier, Future

Generation Computer Systems, vol. 29, Issue: 7, pp. 1645-1660, September 2013. [DOI: 10.1016/j.future.2013.01.010]

[18] M. Aazam, I. Khan, A. A. Alsaffar, E.-N. Huh, "Cloud of Things: Integrating Internet of Things and Cloud Computing and the Issues Involved", in Proceedings of 11th International Bhurban Conference on Applied Sciences & Technology (IBCAST 2014), 14 - 18 January 2014, Islamabad, Pakistan. [DOI: 10.1109/IBCAST.2014.6778179]

[19] M. Aazam, E.-N. Huh, M. St-Hilaire, C.-H. Lung, I. Lambadaris, "Cloud of Things: Integration of IoT with Cloud Computing", Book Chapter, Springer, Robots and Sensor Clouds, pp. 77-94, August 2015.

[20] A. Botta, W. de Donato, V. Persico, A. Pescape, "Integration of Cloud Computing and Internet of Things: a Survey", Journal of Future Generation Computer Systems, vol. 56, pp. 1-54, March 2016. [DOI: 10.1016/j.future.2015.09.021]

[21] M. Aazam, P. P. Hung, E.-N. Huh, "Smart Gateway Based Communication for Cloud of Things", in Proceedings of IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP 2014), 21-24 April 2014, Singapore, Singapore. [DOI: 10.1109/ISSNIP.2014.6827673]

[22] M. Díaz, C. Martin, B. Rubio, "State-of-the-art, challenges, and open issues in the integration of Internet of things and cloud computing", Journal of Network and Computer Applications, vol. 67, No. C, pp. 99-117, May 2016. [DOI: 10.1016/j.jnca.2016.01.010]

[23] M. A. Alsmirat, Y. Jararweh, M. Al-Ayyoub, M. A. Shehab, B. B. Gupta, "Accelerating Compute Intensive Medical Imaging Segmentation Algorithms Using GPUs", Springer, Multimedia Tools and Applications, vol. 76, pp. 3537–3555, September 2016.

[24] B. B. Gupta, D. P. Agrawal, S. Yamaguchi, "Handbook of Research on Modern Cryptographic Solutions for Computer and Cyber Security", IGI Global Publisher, USA, 2016. [ISBN:978-1-5225-0105-3]

[25] Z. Zhang, B. B. Gupta, "Social media security and trustworthiness: Overview and new direction", Elsevier, Future Generation Computer Systems, vol. 28, pp. 914-925, September 2018. [DOI: 10.1016/j.future.2016.10.007]

[26] O. Niggemann, J. R. Kinnebrew, H. Khorasgani, S. Volgmann, A. Bunte, G. Biswas, "Data-Driven Monitoring of Cyber-Physical Systems Leveraging on Big Data and the Internet-of-Things for Diagnosis and Control", in Proceedings

of 26th International Workshop on Principles of Diagnosis, pp. 185-192, 31 August – 3 September 2015, Paris, France.

[27] J. M. Batalla, "Advanced Multimedia Service Provisioning based on efficient interoperability of adaptive streaming protocol and High Efficient Video Coding", Springer, Journal of Real-Time Image Processing, vol. 12, pp. 443-454, March 2015. [DOI: 10.1007/s11554-015-0496-4]

[28] M. Rouse, "IoT security (Internet of Things security)", IoT Agenda, 01/11/2015. [Online]. Available: http://internetofthingsagenda.techtarget.com/definition/IoT-security-Internet-of-Things-security. [Accessed 27/07/2016].

[29] N. Park, N. Kang, "Mutual Authentication Scheme in Secure Internet of Things Technology for Comfortable Lifestyle", MDPI, Sensors vol. 16, no. 1, pp. 1-20, December 2015. [DOI: 10.3390/s16010020]

[30] L. Dong, Z. Han, A. P. Petropulu, H.V. Poor, "Improving Wireless Physical Layer Security via Cooperating Relays", IEEE Transactions on Signal Processing, Vol. 58, Issue: 3, pp. 1875 – 1888, March 2010. [DOI: 10.1109/TSP.2009.2038412]

[31] A. K. Nair, S. Asmi, Dr. A. Gopakumar, "Analysis of Physical layer Security via Co-operative Communication in Internet of Things", Elsevier, Procedia Technology, vol. 24, pp. 896-903, 2016. [DOI: 10.1016/j.protcy.2016.05.162]

[32] W. Hu, H. Tan, P. Corke, W. C. Shih, S. Jha, "Toward trusted wireless sensor networks", ACM Transactions on Sensor Networks, vol. 7, No. 1, Article No. 5, pp. 5-25, August 2010. [DOI: 10.1145/1806895.1806900]

[33] D. Huang, "Mobile cloud computing," IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, vol. 6, no. 10, pp. 27–31, 2011.

[34] Md Whaiduzzaman, M. N. Haque, Md R. K. Chowdhury, A. Gani "A Study on Strategic Provision of Cloud Computing Services", The Scientific World Journal, pp. 1-8, June 2014.

[35] S. K. Garg, S. Versteeg, R. Buyya, "A framework for ranking of cloud computing services", Elsevier, Future Generation Computer Systems, vol. 29, issue: 4, pp. 1012–1023, June 2013. [DOI: 10.1016/j.future.2012.06.006]

[36] G. Skourletopoulos, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, J. N. Sahalos, "An Evaluation of Cloud-Based Mobile Services with Limited Capacity: A Linear Approach", Springer, Soft Computing, vol. 21, issue: 16, pp. 4523-4530, August 2017. [DOI: 10.1007/s00500-016-2083-4]

[37] L. Borovick, R. L. Villars, “The Critical Role of the Network in Big Data Applications”, IDC Analyze the Future, White Paper, Sponsored by: Cisco Systems pp. 1-12, April 2012.

[38] P. Viswanathan, “Cloud Computing – Is it Really All That Beneficial?” abouttech, 07/07/2012. [Online]. Available: http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm. [Accessed 24/01/2015].

[39] F. Pfarr, T. Buckel, A. Winkelmann, “Cloud Computing Data Protection – A Literature Review and Analysis”, in Proceedings of 47th Hawaii International Conference on System Sciences, pp. 5018-5027, 6-9 January 2014, Waikoloa, HI, USA. [DOI: 10.1109/HICSS.2014.616]

[40] E. Almort, S. Andersson, “A study of the advantages & disadvantages of mobile cloud computing versus native environment”, Bachelor Thesis in Software Engineering, School of Computing, Blekinge Institute of Technology, Sweden, May 2013.

[41] S. Fremdt, R. Beck, S. Weber, “Does Cloud Computing Matter? An analysis of the Cloud Model software-as-a-service and its impact on operational agility”, in Proceedings of 46th Hawaii International Conference on System Sciences, pp. 1025-1034, 7-10 January 2013, Wailea, Maui, HI, USA. [DOI: 10.1109/HICSS.2013.182]

[42] Blog: Follow what’s happening at Get Cloud Services, Mobile Cloud Computing – Pros and Cons, [Online], Posted: 23/12/2014, Available: https://www.getcloudservices.com/blog/mobile-cloud-computing-pros-and-cons [Accessed 24/01/2015].

[43] E. Shi, Y. Niu, M. Jakobsson, and R. Chow, “Implicit authentication through learning user behavior”, Springer, in Proceedings of International Conference on Information Security (ISC 2010), Information Security pp 99-11, 25-28 October 2010, Boca Raton, FL, USA. [DOI: 10.1007/978-3-642-18178-8_9]

[44] M. Haghighat, S. Zonouz, M. Abdel-Mottaleb, "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification", Expert Systems with Applications, vol. 42, no. 11, pp. 7905-7916, November 2015.

[45] M. K. Srinivasan, K. Sarukesi, P. Rodrigues, M. S. Manoj, P. Revathy, “State-of-the-art cloud computing security taxonomies: a classification of security challenges in the present cloud computing environment”, in Proceedings of the International Conference on Advances in Computing, Communications and Informatics ICACCI ‘12, pp. 470-476, August 2012. [DOI: 10.1145/2345396.2345474]

[46] B. B. Gupta, O. P. Badve, “Taxonomy of DoS and DDoS Attacks and Desirable Defense Mechanism in a Cloud Computing Environment”, Springer, Neural Computing & Applications, vol. 28, pp. pages3655–3682, 2017.

[48] R. Shaikha, Dr. M. Sasikumar, “Data Classification for achieving Security in cloud computing”, Elsevier, Procedia Computer Science, vol. 45, pp. 493-498, 2015. [DOI: 10.1016/j.procs.2015.03.087]

[49] Y. Kumar, R. Munjal, H. Sharma, “Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures”, IJCSMS International Journal of Computer Science and Management Studies, Vol. 11, Issue: 3, October 2011.

[50] R. Kaur, S. Kinger, "Analysis of Security Algorithms in Cloud Computing", International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 3, pp. 171-176, March 2014.

[51] D. S. Abdul Elminaam, H. M. Abdul Kader, M. M. Hadhoud, “Performance Evaluation of Symmetric Encryption Algorithms”, IJCSNS International Journal of Computer Science and Network Security, vol. 8, No. 12, pp. 280-286, December 2008.

[52] G. Singh, S. Kinger, “Integrating AES, DES, and 3-DES Encryption Algorithms for Enhanced Data Security”, International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July 2013.

[53] A. Sachdev, M. Bhansali, “Enhancing Cloud Computing Security using AES Algorithm”, International Journal of Computer Applications, vol. 9, no. 67, pp. 19-23, April 2013.

[54] U. Somani, K. Lakhani, M. Mundra, “Implementing Digital Signature with RSA Encryption Algorithm to Enhance the Data Security of Cloud in Cloud Computing”, in Proceedings of 1st International Conference on Parallel, Distributed and Grid Computing (PDGC - 2010), 28-30 October 2010, Solan, India.

[55] P. Mell, T. Grance, “The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology”, National Institute of Standards and Technology Special Publication, pp. 1-7, September 2011, [Accessed 24/07/2015].

[56] C. Stergiou, K. E. Psannis, “Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey”, Wiley, International Journal of Network Management, pp. 1-12, May 2016. [DOI:10.1002/nem.1930]

[58] N. Park, M. Kim, H.-C. Bang, “Symmetric Key-Based Authentication and the Session Key Agreement Scheme in IoT Environment”, Springer, Computer Science and its Applications, vol. 330, Conference Paper, pp. 379-384, 2015.

[59] T. Bhattasali, R. Chaki, N. Chaki, “Secure and trusted cloud of thingsin Proceedings of Annual IEEE India Conference (INDICON 2013), pp. 1-6, 13-15 December 2013, Mumbai, India.

[60] Y. Simmhan, A. G. Kumbhare, B. Cao, V. Prasanna, “An analysis of security and privacy issues in smart grid software architectures on clouds”, In Proceedings of IEEE 4th International Conference on Cloud Computing, 4-9 July 2011, pp. 582–589, Washington, DC, USA.

[61] “Synapse Internet of Things Cloud”, Synapse, 2014. https://www.synapse-wireless.com/snap-components/iot. [Accessed 24/07/2015]

[62] N. Grozev, R. Buyya, “Inter-cloud architectures and application brokering: taxonomy and survey”, Wiley Online Library, Software: Practice and Experience, vol. 44, issue: 3, pp. 369–390. March 2014.

[63] K. Jeffery, “ CLOUDs: A large virtualisation of small things’, in Proceedings of the 2nd International Conference on Future Internet of Things and Cloud (FiCloud-2014), 27-29 August 2014, Barcelona, Spain.

[64] B. B. P. Rao, P. Saluia, N. Sharma, A. Mittal, S. V. Sharma, “Cloud computing for Internet of Things & sensing based applications”, In Proceedings of IEEE 6th International Conference on Sensing Technology (ICST 2012), pp. 374–380, 18-21 December 2012, Kolkata, India.

[65] W. He, G. Yan, L. D. Xu, “Developing vehicular data cloud services in the IoT Environment”, IEEE Transactions on Industrial Informatics, vol. 10, Issue: 2, pp. 1587–1595, May 2014.

[66] C. Dobre, F. Xhafa, “Intelligent services for big data science”, Elsevier, Future Generation Computer Systems, vol. 37, pp. 267–281, July 2014.

[67] G. Aceto, A. Botta, W. de Donato, A. Pescape, “Cloud monitoring: A survey”, Elsevier, Computer Networks, vol. 57, issue: 9, pp. 2093–2115, June 2013.

Chapter 5

Security & Privacy of IoE-based Big Data in Cloud Computing

Paper titled "Security & Privacy of IoE-based Big Data in Cloud Computing" authored be C. L. Stergiou and K. E. Psannis to be submitted to IEEE's Internet of Things Journal.

5.1 Introduction

Due to its unique type of services, Cloud Computing could operate as a "base technology" for other technologies. It is a new generation of services that offers the opportunity to the users to access and manage their information, their applications, their data regardless of place and time. Nevertheless, there is a type of service that can include large amounts of data, and it is used to describe the surprisingly rapid increase in the volume of data, which is called Big Data. Both of them faced multiple challenges and issues in their operation. In this work, initially, we illustrate a survey of Cloud Computing and Big Data focusing on security and management challenges of both. Notably, I try to combine the two aforementioned technologies (i.e Big Data & Cloud Computing) to examine their related characteristics, to discover new perspectives and opportunities for their integration. Subsequently, I present how Cloud Computing contributes to Big Data aiming to fill a scientific gap in the field of their integration. Thus, through this analysis, it is shown how Cloud Computing improves the function of Big Data. Finally, I additionally survey the security challenges of the integration of Big Data and Cloud Computing and propose a novel security algorithm. Experimental results presented count on the use of encryption algorithms AES, RC5, RSA, and the proposed model extend the advances of Cloud Computing and Big Data offered a highly novel and scalable service platform to achieve more privacy and security services.

5.2 Related Work

A large number of research works have been contacted the recent years in order to integrate Cloud Computing with Big Data. So, for the purpose of this research we have studied and analyzed previous literature researches which have been studying the integration of Cloud Computing technology with Big Data technology [9] [15-33]. All the works presented here in ascending chronological order.

Starting with, Takabi et al [15] in their work try to explore the roadblocks and solutions aiming to provide a trustworthy Cloud Computing environment.

Agrawal et al [16] introduce a type of tutorial work, which id an organized picture of the challenges that faced by application developers and DBMS (DataBase Management Systems) designers in developing and deploying internet scale applications.

Ji et al [17] introduce several Big Data processing technics from system and application aspects. Specifically, regarding the view of Cloud data management and big data processing mechanisms, Ji et al present the key issues of Big Data processing, including Cloud Computing platform, Cloud architecture, Cloud database and data storage scheme.

Simmhan et al [18] focuses on a scalable software platform for the Smart Grid cyber-physical system using Cloud technologies. Additionally, this platform offers scalable machine-learning models trained over massive datasets for agile demand forecasting, and a portal for visualizing consumption patterns, and validated at the University of Southern California's campus micro-grid.

Talia et al [19], through their work, try to extract useful knowledge from large digital datasets which requires smart and scalable analytics services, programming tools, and applications, by advancing the Cloud from a computation and data management infrastructure to a pervasive and scalable data analytics platform.

In their wok, Demirkan & Delen [20] propose a conceptual framework for DSS in Cloud, and discus about research directions, taking into account a list of requirements for service-oriented DSS which they have defined.

To continue with, Fernandez et al [21] focuses on systems for large-scale analytics based on the MapReduce scheme and Hadoop, and identify several libraries and software projects that have been developed for aiding practitioners in order to address a new programming model.

In another work, Khurana [22] states that BD systems can use different deployment paradigms based on the workloads and access patterns they cater to. Thus, opportunities for tighter integration between CC and BD will enable BD systems to leverage public Cloud environments more effectively.

Moreover, Castelino et al [23] try to present stresses on the integration of BD with CC, which can serve as a driving force for the business and IT industry, as well as, for data analytics in general.

Bohlouli et al [24] proposed a framework that facilitates accessible, efficient and always available knowledge bases for collaborative systems and reduces redundancy and costs by sharing the knowledge between individuals and experts.

In their work, Inukollu et al [25] introduce a discussion about security issues for CC, BD, Map Reduce and Hadoop environment. In particular, the main focus of their work of Inukollu et al is on security issues of CC which are associated with BD.

Ye et al [26] propose a BD driven, Cloud-based information and communication technology (ICT) framework for smart grid. The proposed ICT framework offer price forecast to customers and energy forecast to utility company. Proposed scheme provides confidentiality and nonrepudiation since it performs simultaneously the functions of encryption and digital signature.

Assuncao et al [27] in their work discuss approaches and environments for carrying out analytics on Clouds for BD applications. It revolves around four significant areas of analytics and Big Data. Moreover, Assuncao et al identify possible gaps and offer recommendations for future research directions on Cloud-supported BD computing and analytics solutions.

Hashem et al [28] review the rise of BD in CC, and discuss the relationship between them, BD storage systems, and Hadoop technology.

Yang et al [29] survey BD and CC, and reviews the advantages and the consequences of utilizing CC to tackling BD in the digital earth and relevant science domains.

Dasoriya [30] proposes a scheme for making BD Analytics more accurate, efficient and beneficial. The author realizes that integrations of various tools can be combined to get a more efficient system.

Stergiou and Psannis [31] survey BD and CC and their basic features, focusing on the security and privacy issues of both technologies, and trying to combine the functionality of BD and CC, aiming to examine the frequent features and discover the benefits related in security issues.

Stergiou and Psannis [9] in another work survey BD and CC technologies, additionally with their basic features, focusing on privacy and security challenges, and trying to find out another aspect of combining the functionality of two technologies aiming to examine the benefits related in security challenges of their integration. Summarizing their work, the authors present a new algorithm that can improve CC's security using algorithms providing secure in BD.

Kelbert et al [32] present the "SecureCloud EU Horizon 2020" project, whose goal is to enable new BD applications that use sensitive data in the Cloud without compromising data security and privacy. Thus, the "SecureCloud" designs and develops a layered architecture that allows for three things.

Pargmann et al [33] introduce an approach in which a sound semantical integration of different information types is shown and applied to a concrete use-case, the global monitoring and analysis of wind farms.

Year	Author	Challenge/Issue
2010-Dec.	H. Takabi et al **[15]**	• Authentication & Identity Management • Access Control & Accounting • Trust Management & Policy Integration • Secure-Service Management • Privacy & Data Protection • Organizational Security Management
2011-Mar.	D. Agrawal et al **[16]**	• Single perfect data management solution for Cloud • Extending the Key-Value stores for supporting richer set of applications • Make the systems elastic for effectively utilizing the available resources & minimizing the cost of operation • Designing scalable, elastic, & autonomic multitenant database systems • Ensuring the security & privacy of the data outsourced to the Cloud
2012-Dec.	C. Ji et al **[17]**	• BD Storage & Management • BD Computation & Analysis

		• BD Security
2013-Apr.	H. Demirkan & D. Delen **[20]**	• Need for fast & reliable access to frequently used data by automatically & dynamically partitioning data in-memory across multiple servers, creating continuous data availability & transactional integrity, even in the event of a server failure is very difficult • Local processing power to perform real-time data analysis, in-memory grid computations & parallel transaction & event processing are needed • Virtual runtime environment: runtime control & execution enforcement of ensuring the right work gets done at the right time with the right resources are matched as infrastructure services with client & application sessions based on policy & entitlement
2013-May	D. Talia **[19]**	• Programming abstracts for BD analytics • Data & tool interoperability & openness • Integration of BD analytics frameworks • Data provenance & annotation mechanisms
2013-Aug.	Y. Simmhan et al **[18]**	• Perform intelligent demand-side management & relieve peak load in Smart Power Grids
2014-May	V. N. Inukollu et al **[25]**	• Security issues in CC associated with BD • Challenges of security in Cloud Computing environments can be categorized into network level, user authentication level, data level, and generic issues
2014-Jun.	M. Bohlouli et al **[24]**	• Problem about transferring knowledge about the patient from one hospital to another • Problem in case of incomplete knowledge if patients are to undergo further medication or treatment
2014-Sep.	A. Fernandez et al **[21]**	• BD problems: 1)Scalability, 2)Management, 3)Processing & Analysis
2014-Sep.	A. Khurana **[22]**	• Opportunities for tighter integration between CC & BD will enable BD systems to leverage public Cloud environments more effectively
2014-Oct.	C. Castelino et al **[23]**	• BD performance problem • Analysis & storage problem of BD • Managing, harvesting & analyzing the information obtained from BD
2015-Jan.	I. A. T. Hashem et al **[28]**	• CC & BD integration issues: 1)Scalability, 2)Availability, 3)Data Integrity, 4)Transformation, 5)Data Quality, 6)Heterogeneity, 7)Privacy, 8)Legal/Regulatory Issues, 9)Governance, 10)Data Staging, 11)Distributed Storage Systems, 12)Data Analysis, 13)Data Security
2015-May	M. D. Assuncao et al **[27]**	• BD management challenges: 1)Data Variety, 2)Data Storage, 3)Data Integration, 4)Data Processing & Resource Management
2015-Dec.	F. Ye et al **[26]**	• Price forecast to customers and energy forecast to utility company • Local control centers deal with large amount of data
2017-Jan.	C. Yang et al **[29]**	• BD challenges in CC: 1)Elasticity, 2)Pooled, 3)On-demand, 4)Self-service, 5)Pay-as-you-go • BD technology challenges: 1)Data Storage, 2)Data Transmission, 3)Data Management, 4)Data Processing, 5)Data Analysis, 6)Data Visualization, 7)Data integration,8)Data Architecture, 9)Data Security, 10)Data Privacy, 11)Data Quality
2017-Mar.	F. Kelbert et al **[32]**	• Enable new BD applications that use sensitive data in the Cloud without compromising data security & privacy
2017-Jul.	C. Stergiou & K. E. Psannis **[31]**	• Security and Privacy issues of CC & BD in order to be integrated
2017-Oct.	R. Dasoriya **[30]**	• BD Security & Privacy challenges: 1)Random Distribution, 2)Privacy, 3)Computation, 4)Integrity, 5)Communication, 6)Access Control
2017-Nov.	C. Stergiou & K. E. Psannis **[9]**	• Issues in integration of CC & BD: 1)Data Security, 2)Data Privacy, 3)Data Management
2018-Apr.	H. Pargmann et al **[33]**	• New user-interface style based on augmented reality, allowing an intuitive and quick understanding of complex analysis results

Table 5.1: Related Background Research Work's Challenges and Issues of the Integration Models/Methods of Cloud Computing Technology with Big Data Technology.

5.3 Background Research Analysis

Taking into account the Background & Related Research Section, we realized that the study of finding and achieving an integration model/method of Cloud Computing technology with Big Data technology become more popular in the academic and research community over the recent years. We have come to this conclusion based on our study of several works in this related field. The main balk of these works presented previously, in Background & Related Research Section.

Consequently, the last four years the interest of the researchers has raised considerably compared to previous decade. Figure 5.1 illustrates the growth of studies of finding and achieving an integration model/method of Cloud Computing technology with Big Data technology through the years.

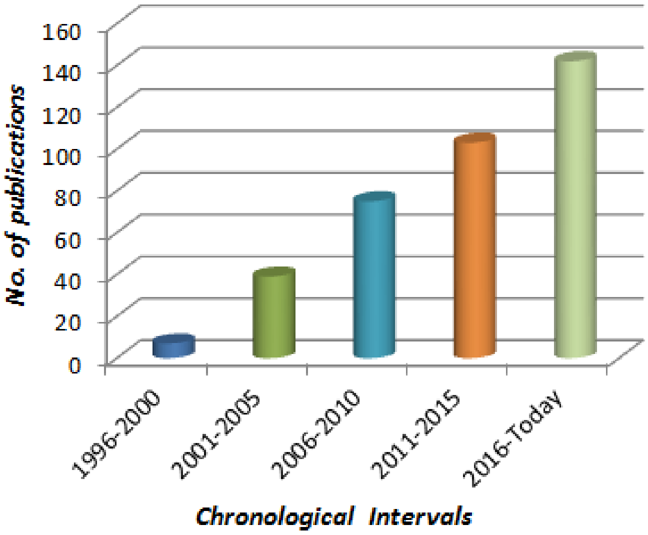


Figure 5.1: Growth of studies of finding and achieving an integration model/method of Cloud Computing technology with Big Data technology.

The study of previous works motivates us to further research the Security and Privacy challenges of the integration of Cloud Computing and Big Data. Thus, based on the main works related to integration model/method between Cloud Computing and Big Data we have tried to demonstrate them in Table 5.1.

Table 5.1 lists the main challenges and issues which we have distinguished from the related background researches. Through the Table 5.1 we can observe which are the major issues and challenges of the integration models/methods of Cloud Computing technology with Big Data technology that have been addressed by the previous works in this field.

Related Background Research	***Challenges***							
	Privacy	*Security*	*Storage*	*Access Control*	*Computation (Processing) & Analysis*	*Management*	*Reliability*	*Scalability*
H. Takabi et al **[15]**	**X**	**X**		**X**		**X**	**X**	
D. Agrawal et al **[16]**	**X**	**X**	**X**		**X**	**X**		**X**
C. Ji et al **[17]**		**X**	**X**		**X**	**X**		
Y. Simmhan et al **[18]**						**X**	**X**	
D. Talia **[19]**					**X**	**X**	**X**	
H. Demirkan & D. Delen **[20]**	**X**		**X**	**X**	**X**		**X**	
A. Fernandez et al **[21]**					**X**	**X**		**X**
A. Khurana **[22]**					**X**	**X**	**X**	
C. Castelino et al **[23]**			**X**		**X**	**X**		
M. Bohlouli et al **[24]**	**X**	**X**			**X**	**X**	**X**	
V. N. Inukollu et al **[25]**		**X**		**X**	**X**	**X**		
F. Ye et al **[26]**					**X**	**X**	**X**	
M. D. Assuncao et al **[27]**			**X**		**X**	**X**		
I. A. T. Hashem et al **[28]**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
C. Yang et al **[29]**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
R. Dasoriya **[30]**	**X**	**X**	**X**	**X**	**X**		**X**	
C. Stergiou & K. E. Psannis **[31]**	**X**	**X**					**X**	
C. Stergiou & K. E. Psannis **[9]**	**X**	**X**				**X**		
F. Kelbert et al **[32]**	**X**	**X**			**X**	**X**		
H. Pargmann et al **[33]**				**X**	**X**	**X**	**X**	

Table 5.2: Background Research challenges in the field of the integration models/methods of Cloud Computing technology with Big Data technology

Table 5.2 presents the challenges which were presented and in most cases have been addresses by the literature work which we have studied. As we could observe, most of the works that we have been studied focus on the “Management” challenge which in our opinion is one of the major issues of Cloud Computing and Big Data integration. As a result the second most popular challenge of the literature work papers is the “Computation (Processing) & Analysis” issue. Additionally, concerning the rest of the challenges we come to the conclusion that the “Reliability” and the “Security” are also basic issues that need to be addressed in both technologies and moreover in their integration model. More detailed, through the number of 20 literature works that we have studied out the statistic results are the following: Privacy 10 of 20, Security 11 of 20, Storage 8 of 20, Access Control 7 of 20, Computation (Processing) & Analysis 16 of 20, Management 17 of 20, Reliability 12 of 20, Scalability 4 of 20. Equally important, as we can observe from the statistical analysis, the less mentioned issues are “Scalability” and “Access Control” which are really vital for the functionality of both technologies. Regarding the statistical analysis of our research, an also count on the recent researches that have been made in the field of Cloud Computing and Big Data, we could realize that more researches need to be done in this field with the aim to find better solutions aiming to improve Security and Privacy issues of integration model of these technologies.

5.4 Cloud Computing

As we already know Cloud Computing could offer some important features to the user such as computing, storage, services, and applications over the Internet [34].

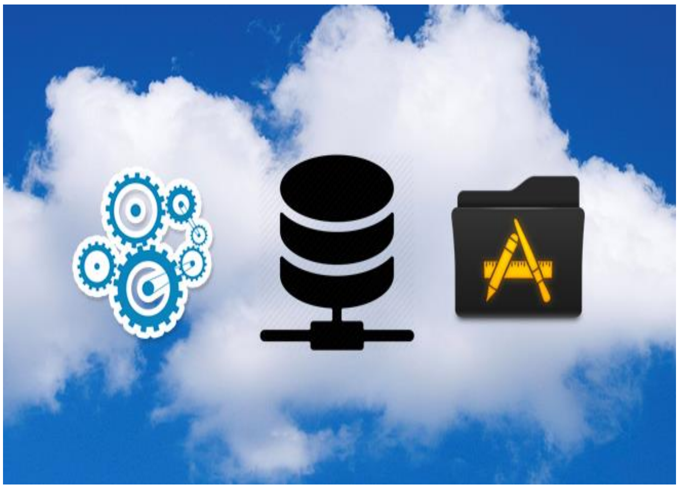


Figure 5.2: Cloud Computing combines Services, Storage & Applications.

Specifically, due to its unique function of Cloud's environment, the Cloud providers and the customers are keen to share the responsibility for security and privacy in Cloud Computing environments; with the limitation however of that the sharing levels will differ for different delivery models, which in turn affect the Cloud extensibility.

More detailed the following are the delivery data models of Cloud Computing [15] [28]:

- *SaaS:* Offers typically enable services by providing a large number of integrated features, which could lead to less extensibility for the customers [15].
- *PaaS:* Aims to enable developers in order to build their own applications on top of the platforms that provided [15].
- *IaaS:* This is the most extensible model. In this model, the Cloud providers must provide some basic, low-level data protection capabilities [15].

5.4.1 Features

As all technologies, so the Cloud Computing technology has some major features which determine its functionality and its 'character'. The major features of Cloud Computing are analyzed and outlined subsequently below.

Storage over Internet

Storage over Internet can be defined as a technology framework that uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks to link servers and storage devices, and to facilitate storage solution deployment. *Storage over*

Internet feature could also be defined as *Storage over Internet Protocol (SoIP)*. The SoIP could offer high-performance and scalable IP storage solutions to the user [35] [36] [37].

Service over Internet

This feature has as main objective to be committed to users in order to help customers aiming to transform aspirations into achievements with the use of Internet's efficiency, speed and ubiquity [35] [36] [37].

Applications over Internet

The feature *Applications over Internet* could be defined regarding to the literature as the programs which can be written to do a job of a current manual task, and which perform their job on the server, such as a Cloud server, via an internet connection [35] [36] [37].

Energy Efficiency

The feature of *Energy Efficiency* could be defined as the way of managing and restraining the increase in energy consumption [35] [36] [37].

Computationally Capable

Cloud Computing could enable services of computational clouds are leveraging the computationally intensive and ubiquitous mobile applications. As a result, a system is considered as *Computationally Capable* when it meets the requirements to offer to the user the expected results, by making the right calculations [35] [36] [37].

5.4.2 Security

The research field of Cloud Computing's security is an emerging and evolving sub-domain of computer's security, network security, and information security. Cloud Computing's security refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of Cloud Computing [37] [38] [39].

Furthermore, the encryption algorithms convert the data into scrambled form by using "*a key*" and only the certified user have the key to decrypt the data [38]. Consequently, count on the literature and the researches that have been made until now the most important encryption technique is the "*Symmetric key Encryption*". In the "*Symmetric key encryption"* only one key is used in order to encrypt and decrypt the data, and the most used algorithm is the *AES* [39] [40].

There are two types of *Symmetric key Encryption* algorithms that are considered the most important for Cloud Computing. Initially, the *AES* (*Advanced Encryption Standard*) encryption algorithm is a novel encryption standard recommended by the NIST aiming to replace former one, DES encryption algorithm. Moreover, it has variable key length of 128, 192, or 256 bits, with default model of 256 bits. In addition, AES encrypts data blocks of 128 bits in 10, 12 and 14 round depending on the key size. AES encryption is fast and flexible, and it can be implemented on various platforms, especially in small devices, such as mobile devices [39] [41].

Algorithm 1 – Sample of AES algorithm

```
    Cipher(byte[] input, byte[] output)
    {
        byte[4,4] State;
        copy input[] into State[] AddRoundKey
        for (round = 1; round < Nr-1; ++round)
        {
            SubBytes          ShiftRows          MixColumns
          AddRoundKey
        }
        SubBytes ShiftRows AddRoundKey
        copy State[] to output[]
}
```

Algorithm 1 represents a sample part of the wide-known AES algorithm. This sample part shows the encryption procedure of AES operation [42].

The other type of *Symmetric key Encryption* algorithm that is considered important for Cloud Computing is the RC5 algorithm. RC5 was designed by Ronald Rivest in 1994. The letters R and C stands for the "*Rivest Cipher*", or alternatively for the "*Ron's Code*". The key length if RC5 has a variable block sizes, 32, 64 or 128 bits, but its normal size is MAX2040 bit, with a block size of 32, 64 or 128. The original suggested choices of parameters were a block size of 64 bits, a 128-bit key and 12 rounds. One of the goals of RC5 was to prompt the study and evaluation of such operations as a cryptographic primitive. Additionally, RC5 consists of a number of modular additions and eXclusive OR, XORs. Regarding its algorithm structure, we can say that it is a Feistel-like network. Thus, the encryption and decryption routines can be specified in a few lines of code. Also, the key schedule, which is in many cases more complex than the other algorithms, expands the key using an essentially one-way function with the binary expansions of both e and the golden ratio as sources of "*nothing up my sleeve numbers*". As a fact, we can observe that the RC5 is basically denoted as *RC5-w/r/b*, where *w=word size in bits*, *r=number of rounds*, *b=number of 8-bit bytes in the key*. Despite all the positives it has, we can say that as a negative is that the speed of this algorithm is slow [39] [43].

Algorithm 2 – Sample of RC5 algorithm

```
    A = A + S[0];
    B = B + S[1];
    for i = 1 to r do
      A = ((A Xor B) <<< B) + S[ 2 * i ]
      B = ((B Xor A) <<< A) + S[ 2 * i + 1]
Next
return A, B
```

Algorithm 2 represents a sample part of the encryption model of RC5 algorithm. This sample part shows the encryption procedure of RC5 operation.

Also, with regard to the other class of encryption algorithms, Asymmetric Key Encryption, the RSA algorithm is very important too. RSA could be could be defined as an internet encryption and authentication system that uses an algorithm developed by three scientists, Ron Rivest, Adi Shamir and Leonard Adleman, in 1977. This algorithm is the most commonly used encryption method. Regarding the usage of RSA, it could be characterized as the only algorithm used for private and public key generation and encryption method. Resulting this, RSA might be the fastest encryption method [39] [44].

Algorithm 3 – Sample of RSA algorithm

```
    Key Generation: KeyGen(p, q)
    Input: Two large primes – p, q
    Compute $n = p . q$
      $\varphi (n) = (p - 1)(q - 1)$
    Choose e such that $gcd(e, \varphi (n)) = 1$
    Determine d such that $e . d \equiv 1 \bmod \varphi (n)$
    Key:
    public key = (e, n)
    secret key= (d, n)
    Encryption:
    $c = m^e \bmod n$
where c is the cipher text and m is the plain text.
```

Algorithm 3 represents a sample part of the wide-known Asymmetric Key Encryption algorithm which called RSA. This sample part shows the encryption procedure of RSA operation. As we can observe from the structure of the RSA algorithm, it has a multiplicative homomorphic property. For example it could easily to find the product of the plain text with just simple mathematical operation, by multiplying the cipher texts. Thus, the result of this operation would be the cipher text of the product. More specifically:

Given $c_i = E(m_i) = m_i^e \bmod n$, then

$$\mathbf{(c1 . c2) \bmod n = (m_1 . m_2)^e \bmod n}$$

5.4.3 Privacy & Security challenges in Cloud Computing

As we already know from the literature review, the Cloud Computing environments could be characterized as multi-domain environments. These environments could treat each domain as it can use different security, privacy, and trust requirements, and could lead them to employ various mechanisms, interfaces, and semantics. Thus, we realize that those Cloud domains might represent separate enabled services or other infrastructural or application components [15].

As a result, and regarding the literature review, we can conclude that the major challenges in the field of Cloud Computing are the following:

- ✓ Authentication and Identity Management
- ✓ Access Control and Accounting

- ✓ Trust Management and Policy Integration
- ✓ Secure-Service Management
- ✓ Privacy and Data Protection
- ✓ Organizational Security Management

We can realize that the main concept of these challenges focus on the *securing the managing services of the data* in a Cloud Computing environment.

5.5 Big Data

Nowadays, there are tremendous amounts of data which generated in daily base in the sectors of manufacturing, business, science and peoples' personal lives. As a result, we can conclude that the proper processing of the data could reveal new knowledge about market, society and environment, additionally with enabling people to react emerging opportunities and changes in a timely manner [45] [46].

Furthermore, the accustomed data processing technologies, such as database sets and data warehouse, are becoming inadequate to the huge amounts of data that is needed to deal with. These challenges are known as Big Data and consists a novel field of study for the current researchers. Thus, due to its importance and commonness, it has gained enormous attention recently [45] [46].

5.5.1 Big Data's characteristics

Big Data defined as "*a big thing in the field of modern technologies*" [46]. To better understand the 'meaning' of Big Data, we could have to find out the usage of their five major characteristics, which are widely known as five Vs of Big Data [47].

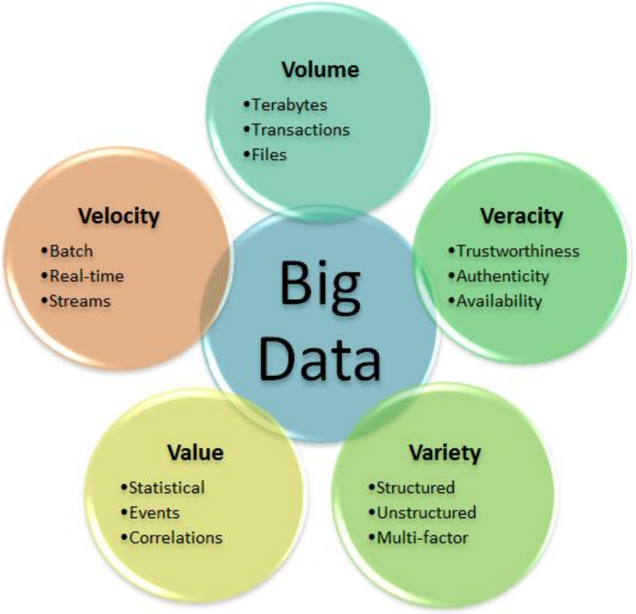


Figure 5.3: 5Vs of Big Data and their major characteristics.

The five Vs of Big Data are:

- **Volume:** vast quantity of data which are generated each second.
- **Velocity:** speed at which the new data sets are generated and additionally speed at which the data sets move around.
- **Variety**: various types of data that can be used.
- **Veracity**: messiness or trustworthiness of the data.
- **Value:** worth of the data which have being extracted.

5.5.2 Big Data issues and challenges

Regarding our research we came to the conclusion that the major issues of Big Data that is fundamental needed to be addressed are three. These are: storage, management, and processing. Each of these challenges represents a large set of technical research challenges in its own right.

BDC1: Big Data Storage

Data's quantity has exploded each time a new storage medium was invented. Also, data creation does not have any restriction; it could be created by everyone and everything (e.g., devices, etc), and it is not just, as heretofore, by professionals such as scientist, journalists, writers, etc. [48].

BDC2: Big Data Management

Big Data management could be used with focus on customize the consistency level. As a useful aspect of the applications could be considered the customized replication and consistency enforcements, where a number of updates may require higher integrity and some may require the higher scalability of relaxed consistency [49].

Moreover, the *HBase* is a significant implementation of NoSQL model in the Hadoop project. *Hbase* is a distributed column-oriented database which was built on top of HDFS (Hadoop Distributed File System. *HBase* does not support SQL model due it is not a relational database. However, this system has the ability to host very large, sparsely populated tables on clusters made from commodity hardware. In this system the data stored in rows and column family group rows. In *HBase*, tables are partitioned horizontally into regions, which are the units that get distributed over the *HBase* cluster [49].

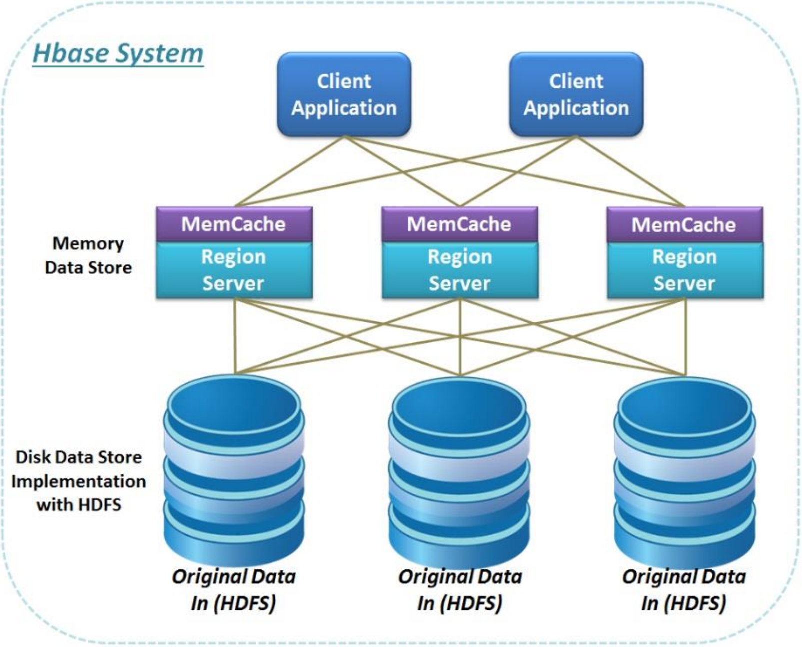


Figure 5.4: Architecture of the *HBase* NoSQL Database System management.

Count on previous researches, such as [50], where noted that "*there is no universally accepted way to store raw data, … reduced data, and … the code and parameter choices that produced the data*", data and the source of the information of the data become a critical challenge. Moreover, count on the same former research [50], the authors also notes that "*We are unaware of any robust, open source, platform-independent solution to this problem*". Thus, we now could easily realize that this fact is still remains. Resulting our research, we can reach to the agreement that there is not established yet such a perfect way of Big Data management.

BDC3: Big Data Processing

Let's suppose that an exabyte of data needs to be processed in its wholeness. More simply, we can assume the data is crumbled into blocks of 8 words, and as a result an exabyte is equal to *1 Kilo* petabytes. Considering that a processor could expend 100 instructions on one block at 5 gigahertz, and then the time which required for end-to-end processing would be about 20 nanoseconds.

5.5.3 Big Data Security & Compliance

In some areas, for example the social media and the health information, as more data is gathered for each person, there is a fear that some organizations would extract much information about people. One critical scenario could be the data collected in the electronic health record systems count on the HIPAA/HITECH provisions which is already increasing concerns about violations of one's privacy. A rough and 'easy' solution is to develop algorithms that have the ability to randomize personal data among a large data set in order to ensure privacy [48].

Regarding the current status of data, maybe the major issue about it could be the unregulated accumulation of data by numerous social media companies, such as

Facebook, Instagram etc. This type of data could be considered as a strict security and privacy concern due to the fact that a huge number of people are willing to give without a second thought much personal information [48].

Based on the related work we have studied, the International Data Corporation (IDC) invented the term "*digital shadow*" with the aim to reflect the amount of data concerning to a person from who has been collected, organized, and analyzed, aiming to create an aggregate "*picture*" of this person. The main problem that arises is the fact that how much of this information that have been created, ignoring whether it is true or false, could be remain private [48].

5.5.4 Big Data Sources

All the data which is related to the term Big Data have a specific origin. This origin, or we can call it better as source, could give various types of data. This fact could be based on one of the 5Vs of Big Data, the Variety. The sources of the Big Data could drive us in some challenges that need to be addressed in the overall use of Big Data. Particular, we will try to address a number of major Big Data sources and some challenges that arise from them.

<u>BDS1: Earth Sciences</u>

Collection and generation of large data sets in every second and at different space-time scales for operations as presenting, monitoring and understanding complex earth systems are enabled by the preferment of sensing and computing simulation technologies. So, as an example Earth Observation software collects terabytes of images on a daily bases [29] [51] with a gradual increase of space, time and spectral analysis [29] [52] [53] [54].

<u>BDS2: Internet of Things</u>

The Internet of Things (IoT) consist of "*a network of physical objects, devices, vehicles, buildings and other items that have been embedded with electronics, software, sensors, and network connectivity, and also have the ability to permit these objects to gather and interchange data*" [29] [55] [56] [57] [58] [59]. The whole data that can be generated from the various IoT sensors encloses spatiotemporal information, and thus it can describe as Big Data. The combined use of IoT and Big Data in network environments, and in addition integrated with technologies such as Cloud Computing, could offer new opportunities and could lead us in the acceleration developing of Smart Cities [29] [60] [61] [62].

<u>BDS3: Social Sciences</u>

Big Data could also be generated by the various social networks, like Instagram, Twitter and Facebook, and thus they could transform social sciences [29] [63]. Due to this fact, Big Data mining methods used by sociologists, political scientists, economists and other social scholars in order to analyze the various social interactions, the health records, the phone logs, the government records and other digital traces such as these [29] [64].

BDS4: Business

The various decisions for strategy, managing optimization and competition related to Big Data could be enhanced by *business intelligence and analytics* [29] [65] [66]. Data related to the previous scenarios contains harmful amounts of geospatial information, for example where and when a transition occurred [67] [68] [69] [70].

BDS5: Industry

In the software Industry 4.0, which is a fourth industrial revolution, the products and production systems leverage technologies such as IoT and Big Data aiming to build ad-hoc networks for self-control and self-optimization [71] [72] [73] [74].

5.6 Cloud Computing & Big Data Integration

In order to achieve the best integration model between Cloud Computing and Big Data several researches need to be done. Count on previous related works and count on our research we could introduce the contributions of the features of Cloud Computing in Big Data.

Cloud Computing Features **Big Data Features (5 Vs)**	*Storage over Internet*	*Service over Internet*	*Applications over Internet*	*Energy efficiency*	*Computationally capable*
Volume	**X**		**X**	**X**	
Velocity		**X**	**X**	**X**	**X**
Variety		**X**		**X**	
Veracity	**X**		**X**		**X**
Value	**X**	**X**	**X**	**X**	**X**

Table 5.3: Contribution of Big Data in Cloud Computing

Table 5.3 lists the features of Cloud Computing technology regarding the convenience this technology provides. Also, it enumerates the main features, also known as 5 VS, of the Big Data. The main purpose of Table 5.3 is to show which of the specific features of Cloud Computing technology, contribute more and as a result related more to the features of Big Data technology. As we can observe from Table 5.3, the feature of Big Data which affected more by the features of Cloud Computing is "*Value*". Value as we have already noticed previously relates to the data that consists large data sets. Thus, it should obviously the most significant feature of Big Data contributing with the technology of Cloud Computing. As regards the Cloud Computing, the features which affected more are "*Applications over Internet*" and "*Energy Efficiency*". Both, these two features based on the use of the data through the network. Applications could extract large amounts of data sets and also the grouping of data in large data sets can lead to better use of the power resources which leads to a more energy efficient environment. As a general conclusion, we can observe that those two technologies contribute more each other in many of their features.

Cloud Computing Models	***SaaS*** *Software as a Service*	***PaaS*** *Platform as a Service*	***IaaS*** *Infrastructure as a Service*
Big Data Sources (*BDS*)			
Earth Sciences (***BDS1***)	**X**		**X**
Internet of Things (***BDS2***)	**X**	**X**	**X**
Social Sciences (***BDS3***)	**X**	**X**	
Business (***BDS4***)	**X**		**X**
Industry (***BDS5***)		**X**	**X**

Table 5.4: Contribution of Cloud Computing Models in Big Data Sources

Table 5.4 lists the three models of Cloud Computing technology and the basic sources of Big Data. Through Table 5.4 we can relate the necessity of using the different models of Cloud Computing bottles to the various sources that export Big Data. As we can observe from Table 5.4, the source of Big Data which contributed more by the models of Cloud Computing is "*Internet of Things*". Internet of Things, regarding our research, is the major source that Big Data counts on. Additionally, due to its use which is especially connected to the internet makes it quite close to all the models of Cloud Computing technology. Regarding the Cloud Computing, the models which contributed more are "*SaaS*" and "*IaaS*". So, we can be led to the conclusion that Cloud Computing can contribute to Big Data providing both software and hardware resources in order to be produced large data sets.

Through the integration of Big Data and Cloud Computing we have the opportunity to expand the use, the management, and the transmission of the large data sets which constitute the Big Data, provided in environments based on Cloud Computing. Applications and information that could be produced by Big Data, with this integration can be used through the Cloud storage. The integration of Big Data and Cloud technologies represented in Figure 5.5. Cloud Computing offers to mobile and wireless users to access all the information and the application that needed for all the data that can defined as Big Data.

Figure 5.5: Big Data & Cloud Computing Integration.

5.6.1 Challenges and issues in Big Data and Cloud Computing integration

The virtually resources and unlimited capabilities of Cloud Computing aiming to balance its technological constrains, such as storage , communication and processing, could offer beneficial use to Big Data. Furthermore, Big Data could offer a dynamic manner by delivering novel services in the world extracting the meaning of the large scale data sets which make up it, take into advantage the benefits offered by the Cloud Computing. In many cases, Cloud Computing could offer the intermediate layer between the data and the applications, hiding all the functionalities and complexity that would be necessary to be implemented.

Big Data & Cloud Computing Integration Challenges	*Privacy*	*Security*	*Storage*	*Access Control*	*Computation (Processing) & Analysis*	*Management*	*Reliability*	*Scalability*
Big Data		**X**	**X**		**X**	**X**		**X**
Cloud Computing	**X**	**X**	**X**	**X**		**X**	**X**	

Table 5.5: Integration Challenges Effects of Big Data & Cloud Computing

Table 5.5 shows the Cloud Computing and Big Data integration challenges which were presented and in most cases have been addresses by the literature work which we have studied. With this table we can show the contribution of Cloud Computing and Big Data integration challenges on its technology separately. More specifically, the Table 5.5 reveals that the common challenges that affected of both technologies are "*Security*", "*Storage*", and "*Management*". Take into account the conclusion drawn from Table 5.2, with which showing that most of the works that we have been studied focus on the "*Management*" challenge, so we can to strengthen our view that this is the most important challenge of the integration of Cloud Computing and Big Data. Additionally, the other two important challenges resulting could also lead us to the conclusion that the store of the large data sets, and then to make those data sets more secure play a vital role to the integration of Cloud Computing and Big Data.

We have studded multiple works that addresses a number of significant challenges and problems that pertaining to two procedures, which are the storage and the processing of Big Data in the Cloud. Nowadays, there are a small number of tools which are available to the users in order to address the multiple challenges of Big Data processing in Cloud environments. Also, novel technologies and techniques related to many important Big Data applications are not able to solve the actual problem of storing and querying Big Data [28].

Data staging

Due to the literature research, the most important open research issue as regards the data staging is related to the heterogeneous nature of data. As we know, there will be challenging tasks by transforming and cleaning unstructured data like this before loading them into the storage for analysis. Nevertheless, when meaningful information required, we have to understand the context of unstructured data which is necessary [28].

Distributed storage systems

Through the recent years, researchers have proposed various solutions in order to store and retrieve large amounts of data. Nevertheless, several issues prevent the successful implementation of these solutions, including the capability of current Cloud technologies, aiming to provide necessary capacity and high performance, in order to address massive quantity of data [75], optimization of existing file systems for the volumes required by data mining applications [28].

Data analysis

Regarding the Data Analysis, selection of an appropriate model for large amounts of data analysis is critical. One important work in this field introduced by Talia [76], pointed out that obtaining useful information from huge quantity of data requires scalable analysis algorithms aiming to produce timely results. Consequently, in order to process data such this required efficient data analysis tools and techniques [28] [77].

Data security

Security threats in Cloud environment are magnified by the variety, volume and velocity of Big Data. Moreover, a various number of threats and problems, such as integrity, confidentiality, availability and privacy of data, exist in Big Data using Cloud Computing platforms. Consequently, in order to be more secured, Cloud vendors must ensure that all service level agreements are complied with [28].

5.6.2 Security challenges in Big Data and Cloud Computing integration

There is a prompt and independent evolution considering the two words of Big Data and Cloud Computing. Also, Big Data technology extends its scope to deal with various types of data in a more distributed and dynamic manner and by offering new techniques in various aspects of real life scenarios, might be benefit from the use of Cloud Computing. Frequently, Cloud could offer the intermediate layer between the users and the systems that the large amounts of data exist, hiding all the complexity and functionalities [78] [79].

Big Data	*Big Data Storage (**BDC1**)*	*Big Data Management (**BDC2**)*	*Big Data Processing (**BDC3**)*
Cloud Computing			
Authentication & Identity Management		**X**	**X**
Access Control & Accounting	**X**		**X**
Trust Management & Policy Integration		**X**	

Secure-Service Management	X	X	
Privacy & Data Protection	X		
Organizational Security Management	X	X	X

Table 5.6: Big Data & Cloud Computing Security Challenges Affections

Table 5.6 lists the security challenges that both Cloud Computing and Big Data face. As already analyzed above, in Sections 5.4 and 5.5, the major security challenges of Big Data, regarding the literature are three, Big Data Storage, Big Data Management and Big Data Processing, and from the other hand the major security challenges of Cloud Computing, regarding the literature are six, Authentication & Identity Management, Access Control & Accounting, Trust Management & Policy Integration, Secure-Service Management, Privacy & Data Protection and Organizational Security Management. As we can observe Organizational Security Management is related to both three Big Data Security challenges. Thus, we easily be able to understand that the words “Security” and “Management” play an important role in the life of Cloud Computing and Big Data both individually and in their integrated form.

Through the integration of Cloud Computing and Big Data could be shown that Cloud Computing can fill some gaps of Big Data such the storage and management. Furthermore, the security challenge of the integration of Cloud Computing and Big Data is a serious aspect.

5.6.3 Proposed Security Method for Big Data Encryption in Cloud Environment

Through the literature research, we reach in the fact that AES, RC5 and RSA are the fastest and more efficient encryption algorithms. Moreover, AES also considered providing a better security environment for Cloud Computing technology, counting on previous works that have been done in this field. Furthermore, the symmetric encryption method of RC5 and the asymmetric encryption method of RSA give the opportunity to the users to achieve a more secure Cloud environment.

Algorithms	**AES**	**RC5**	**RSA**
Characteristics			
Year Founded	1998	1994	1977
Key Length	128, 192, or 256 bits	32, 64, or 128 bits	1024 – 4096 bits
Rounds of Encryption	10, 12, or 14	12	1
Type of Key Cryptography	Symmetric Key Encryption	Symmetric Key Encryption	Asymmetric Key Encryption
Certification	AES winner, CRYPT, REC, NESSIE, NSA	AES finalist	PKCS#1, ANSI X9.31, IEEE 1362
Algorithm Speed	Very Fast	Fast	Very Fast

Table 5.7: Comparison of AES, RC5 and RSA Algorithms

Table 5.7 lists the key characteristics of the three aforementioned encryption algorithms (AES, RC5, RSA) which have been studied and used with the aim to reach our experimental proposal. As we have already mentioned before, one of their key features which is important in our study is their speed. Both three algorithms could be characterized as fast algorithms, due to their quick response in their encryption procedure. Therefore, one key characteristic in which they differ a lot is the Rounds of their encryption method, where AES needs 10, 12, or 14 rounds, and RC5 needs 12 rounds, instead of RSA which needs only 1 round.

Our proposed method collects and combines all the benefits of the three algorithms in order to provide a better encryption scenario for Big Data in Cloud Computing environments. Thus, with our proposed model we can amplify the advances of Cloud Computing and Big Data technologies, by offering a highly novel and scalable efficient service platform. As a result, we can propose and introduce the following algorithm for a more secure use of Big Data in Cloud Computing.

Algorithm 4 – Proposed algorithm

```
Key Production Procedure
m1 = im/8
counter = 0
while m1
   if (m1==‘ ’ or m1 ==‘space’)
      break
   else
      counter = counter + 1
k = im * counter
------------------------------------------
Encryption Procedure
input k
input im
kc = 0
wpc = 0
while k
   if (k==‘ ’ or k ==‘space’)
      break
   else
      kc = kc + 1
while im
   if (im ==‘ ’ or im ==‘space’)
      break
   else
      wpc = wpc + 1
om = 1
while kc > 0
   for i=1, i++, i<=wpc
      om = (om*i)+im
nk = k + i
transfer routine for nk & om
```

m1 = key production number	**wpc** = word package counter
im = input package of data	**om** = output package of data
counter = word counter	**i** = default counter
k = key	**nk** = new encryption key
kc = key word counter	

Table 5.8: Contents of the Algorithm

5.7 Comparative Analysis

In addition to the literature review of past works presented in Section 5.2, we also review some major works in the the specific field of Cloud Computing and Big Data integration, with the aim of security. In this section we will make a comparative analysis study of these previous works which we have distinguished in the following paragraphs. Initially, we analyze what each of them deals with, presented from the oldest to the most recent.

K. Gai et al [80] concentrate on privacy issue and propose a new data encryption approach, which is called *Dynamic Data Encryption Strategy* (D2ES), which focuses to selectively encrypt data and use privacy classification methods under timing constraints. Finally, they present experiments that evaluate the performance of the proposed D2ES, which provide the proof of the privacy enhancement.

H. Matallah et al [81] propose an approach in order to improve the service metadata for Hadoop to maintain consistency without much compromising performance and scalability of metadata by suggesting a mixed solution between centralization and distribution of metadata for enhancing the performance and scalability of the model.

S. K. Mishra et al [82] examined the energy consumption in Cloud Computing environment count on varieties of services and achieved the provisions that promoting Green Cloud Computing. Also, the authors proposed an adaptive task allocation algorithm for the heterogeneous Cloud environment, which is trying to minimize the makespan of the cloud system and reduce the energy consumption. Furthermore, the proposed evaluation scenarios tested in CloudSim simulation environment.

R. Chaudhary et al [83] propose an innovative SDN-based Big Data management scenario with respect to the optimized network resource consumption such as network bandwidth and data storage units. Moreover, with the use of their proposed solution, they believe that the developers can deploy and analyze real-time traffic behavior for the future Big Data applications in MCE.

Y. Wen et al [84] model the problem of how to schedule workflow with such data privacy protection constraints, while minimizing both execution time and monetary cost for Big Data applications on Cloud environment as a multi-objective optimization problem and propose a Multi-Objective Privacy-Aware workflow scheduling algorithm, named MOPA. This proposed algorithm can offer to cloud customers a set of Pareto tradeoff solutions.

Big Data & Cloud Computing Integration Challenges	*Privacy*	*Security*	*Energy Efficient*	*Access Control*	*Computation (Processing) & Analysis*	*Management*
Gai et al. [80]	**X**	**X**				
Matallah et al. [81]					**X**	**X**
Mishra et al. [82]			**X**		**X**	
Ghaudhary et al. [83]				**X**	**X**	**X**
Wen et al. [84]	**X**	**X**				
Proposed method	**X**	**X**		**X**	**X**	**X**

Table 5.9: Related Work Comparison

Table 5.9 shows that most of the relative works are involving and trying to improve challenges related to *Computation (Processing) & Analysis*. Furthermore, the most of the former relative works deal with the *Privacy*, *Security*, and *Management*. The challenges that are not in the top of the research interest are *Energy Efficiency* and *Access Control*. Consequently, we could reach the conclusion that there are many open issues in the field of Cloud Computing and Big Data integration that need to be solved. To sum up, we realize the need of more research in the area of Security and Management of Big Data in Cloud Computing environment and thus we propose a new encryption method of the data in a Cloud environment which will try to deal and improve challenges such as Management and Security of data, that in many cases are Big Data.

5.8 Experimental Analysis & Results

5.8.1 Experimental Comparative Analysis

Considering the benefits of the security models and algorithms of Cloud Computing and Big Data we could have a beneficial use of integration model. In addition, the novel integration model could has good potential for benefiting from instruction-level parallelism and will support any type of block sizes and key sizes used mostly for Big Data in Cloud Computing environment.

Through this integration a number of useful operations could be achieved, i.e. we can use the Cloud-based services and applications with the aim to connect devices and sensors, to produce large amounts of data, and also made them capable to share their data through the Cloud, by reducing the security challenges.

Nevertheless, many other issues and benefits need to be addressed through the integration of Cloud Computing and Big Data, regarding the security and management challenges, but in addition regarding the whole use of both technologies together.

AES Characteristics	Key length	Rounds	Certifications	Speed
Cloud Computing	X	X	X	
Big Data	X		X	X
Integration Model	X	X	X	X

Table 5.10: AES Contribution in Cloud Computing and Big Data.

RC5 Characteristics	Key length	Rounds	Certifications	Speed
Cloud Computing	X	X	X	
Big Data	X			X
Integration Model	X	X	X	X

Table 5.11: RC5 Contribution in Cloud Computing and Big Data.

RSA Characteristics	Key length	Rounds	Certifications	Speed
Cloud Computing	X	X	X	
Big Data			X	X
Integration Model	X	X	X	X

Table 5.12: RSA Contribution in Cloud Computing and Big Data.

The three Tables above (Table 5.10, Table 5.11 and Table 5.12) revealing the key characteristics of the three encryption algorithms discussed previously, and used aiming to achieve an integration of the Cloud Computing and Big Data concerning the security challenges, and a little bit the management of the data challenges. Table 5.10 demonstrates which of the major features of AES encryption algorithm contributes both Cloud Computing and Big Data, and finally how completely contributes their integration model. Additionally, Table 5.11 demonstrates which of the major features of RC5 encryption algorithm also contributes both Cloud Computing and Big Data, and finally how completely contributes their integration model. At the end, Table 5.12 demonstrates which of the major features of RSA encryption algorithm also contributes both Cloud Computing and Big Data, and finally how completely contributes their integration model too.

5.8.2 Experimental Comparative Results

In order to prove the beneficial operation of our proposed model a number of simulations have been made. With the experimental scenarios which we have made we compared in a period the function of our proposed model compared with existed algorithms of AES, RC5 and RSA. Through these simulations and their results we have the opportunity to realize that a good effort have been done in order to offer a more secure and efficient model.

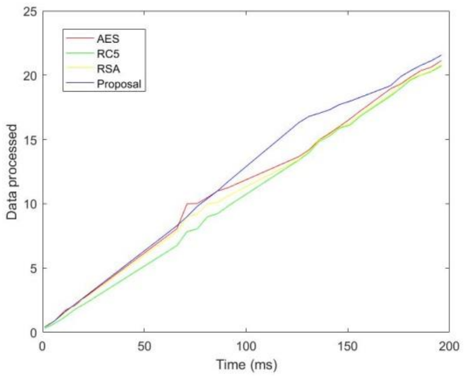


Figure 5.6: Performance of data processed comparison of the AES, RC5, RSA & Proposed model.

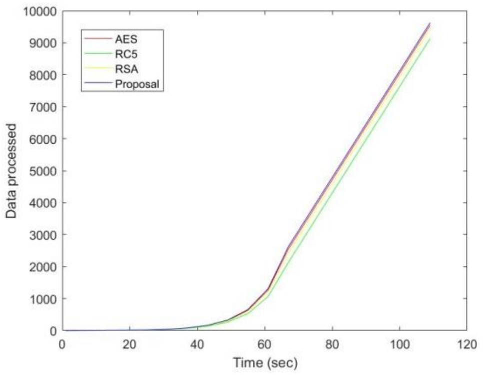


Figure 5.7: Performance of data processed comparison of the AES, RC5, RSA & Proposed model.

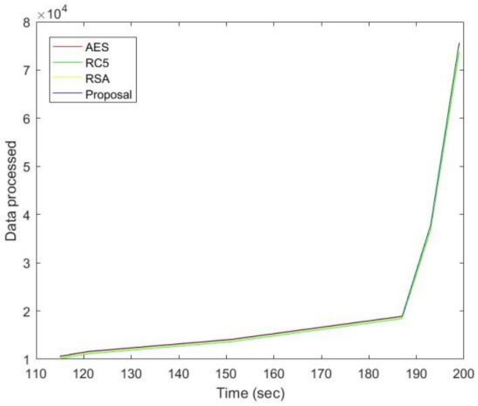


Figure 5.8: Performance of data processed comparison of the AES, RC5, RSA & Proposed model.

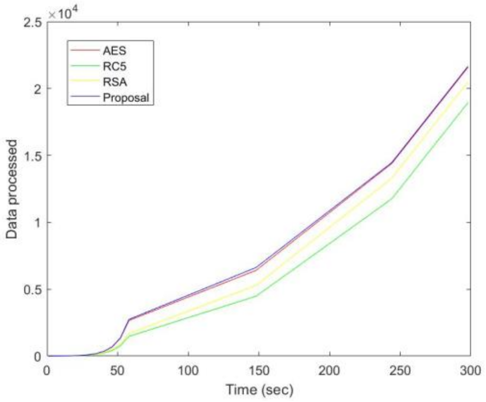


Figure 5.9: Performance of data processed comparison of the AES, RC5, RSA & Proposed model.

Figures 5.6, 5.7, 5.8, and 5.9 show four experimental scenarios that considering the performance of data processing, with the use of the four algorithms (AES, RC5, RSA & Proposed model) in the measure of time. Through these scenarios we can observe that by applying our proposed model we could achieve better data processed in comparison with the existed encryption algorithms AES, RC5 and RSA. Also, in figures 5.6, 5.7, 5.8, and 5.9 the following clarifications are given: line red represents AES, line green represents RC5, line yellow represents RSA, and line blue represents our proposed model. Thus, as we can observe from figures 5.6, 5.7, 5.8, and 5.9 in the same time with our proposed model we could process larger amount of data.

5.9 Chapter Summary

It is wide known that Cloud Computing provides many possibilities, but also places several limitations as well. Thus, in this work we presented a study of Cloud Computing and Big Data aiming to the security and management challenges of them. Particularly, we have combined them in order to testify the related characteristics, and in order to find out the benefits of their integration. Consequently, we presented the contribution of Big Data to the Cloud Computing aiming to fill the existed scientific gap on this field. Moreover, this work showed how Cloud Computing improves the function of Big Data. Finally, we surveyed the security issues of Cloud Computing and Big Data integration and propose a novel security model. Experimental results also presented in the end of this work count on the use of encryption algorithms AES, RC5, RSA and the proposed model extend the advances of Cloud Computing and Big Data offered a highly novel and scalable efficient service platform in order to achieve more privacy and secure services.

Finally, security problems of Cloud Computing and Big Data integration were studied through the proposed algorithm model, and in addition to this it shown how the three encryption algorithms analyzed in this work contributes in the integration model of Cloud Computing and Big Data. Consequently, count on this we could set the field of further research in the future, on the integration of Cloud Computing and Big Data. Count on the hasty development of them, the security and management challenges of the data streamed and stored in the Cloud must be clarified or minimized, with the aim to have an efficient integration model. These security issues that studied in this work could be the field for future research as a case study, with the goal of minimizing them.

5.10 Chapter References

[9] C. Stergiou, K. E. Psannis, "Efficient and secure BIG data delivery in Cloud Computing", Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017.

[15] H. Takabi, J. B. D. Joshi, G.-J. Ahn, "Security and Privacy Challenges in Cloud Computing Environments", IEEE Security & Privacy, vol. 8, issue: 6, pp. 24-31, December 2010. [DOI: 10.1109/MSP.2010.186]

[16] D. Agrawal, S. Das, A. El Abbadi, "Big data and cloud computing: current state and future opportunities", ACM, in Proceedings of the 14th International Conference on Extending Database Technology, EDBT/ICDT '11, pp. 530-533, 21-24 March 2011, Uppsala, Sweden.

[17] C. Ji, Y. Li, W. Qiu, U. Awada, K. Li, "Big Data Processing in Cloud Computing Environments", in Proceedings of IEEE 12th International Symposium on Pervasive Systems, Algorithms and Networks, 13-15 December 2012, San Marcos, TX, USA.

[18] Y. Simmhan, S. Aman, A. Kumbhare, R. Liu, S. Stevens, Q. Zhou, V. Prasanna, "Cloud-Based Software Platform for Big Data Analytics in Smart Grids", IEEE Journal of Computing in Science & Engineering, vol. 15, issue: 4, pp. 38-47, August 2013.

[19] D. Talia, "Clouds for Scalable Big Data Analytics", IEEE Computer, vol. 46, issue: 5, pp. 98-101, May 2013 [DOI: 10.1109/MC.2013.162]

[20] H. Demirkan, D. Delen, “Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud”, Elsevier, Decision Support Systems, vol. 55, issue: 1, pp. 412-421, April 2013. [DOI: 10.1016/j.dss.2012.05.048]

[21] A. Fernandez, S. del Rio, V. Lopez, A. Bawakid, M. J. del Jesus, J. M. Benitez, F. Herrera, “Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks”, Wiley Online Library, Wires Data Mining and Knowledge Discovery, vol. 4, issue: 5, pp. 380-409, September 2014.

[22] A. Khurana, “Bringing Big Data Systems to the Cloud”, IEEE Cloud Computing, vol. 1, issue: 3, pp. 72-75, September 2014. [DOI: 10.1109/MCC.2014.47]

[23] C. Castelino, D. Grandhi, H. G. Narula, N. H. Chokshi, “Integration of Big Data and Cloud Computing”, International Journal of Engineering Trends and Technology (IJETT), vol. 16, no. 2, pp. 100-102, October 2014.

[24] M. Bohlouli, F. Merges, M. Fathi, “Knowledge integration of distributed enterprises using cloud based big data analytics”, in Proceedings of IEEE International Conference on Electro/Information Technology, 5-7 June 2014, Milwaukee, WI, USA.

[25] V. N. Inukollu, S. Arsi, S. R. Ravuri, “Security Issues Associated with Big Data in Cloud Computing”, International Journal of Network Security & Its Applications (IJNSA), Vol.6, No.3, pp. 45-56, May 2014.

[26] F. Ye, Y. Qian, R. Q. Hu, “An Identity-Based Security Scheme for a Big Data Driven Cloud Computing Framework in Smart Grid”, in Proceedings of IEEE Global Communications Conference (GLOBECOM 2015), 6-10 December 2015, San Diego, CA, USA.

[27] M. D. Assuncao, R. N. Calheiros, S. Bianchi, M. A. S. Netto, R. Buyya, “Big Data computing and clouds: Trends and future directions”, Elsevier, Journal of Parallel and Distributed Computing, volumes 79-80, pp. 3-15, May 2015.

[28] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, S. U. Khan, “The rise of “big data” on cloud computing: Review and open research issues”, Elsevier, Information Systems, vol. 47, pp. 98-115, January 2015.

[29] C. Yang, Q. Huang, Z. Li, K. Liu, F. Hu, “Big Data and cloud computing: innovation opportunities and challenges”, International Journal of Digital Earth, vol. 10, issue: 1, pp. 13-53, 2017. [DOI: 10.1080/17538947.2016.1239771]

[30] R. Dasoriya, “A review of big data analytics over cloud”, in Proceedings of IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia 2017), 5-7 October 2017, Bangalore, India.

[31] C. Stergiou, K. E. Psannis, “Algorithms for Big Data in Advanced Communication Systems and Cloud Computing”, in Proceedings of IEEE 19th Conference on Business Informatics (CBI 2017), 24-27 July 2017, Thessaloniki, Greece.

[32] F. Kelbert, F. Gregor, R. Pires, S. Kopsell, M. Pasin, A. Havet, V. Schiavoni, P. Felber, C. Fetzer, P. Pietzuch, “SecureCloud: Secure big data processing in untrusted clouds”, in Proceedings of Design, Automation & Test in Europe Conference & Exhibition (DATE 2017), 27-31 March 2017, Lausanne, Switzerland.

[33] H. Pargmann, D. Euhausen, R. Faber, “Intelligent big data processing for wind farm monitoring and analysis based on cloud-technologies and digital twins: A quantitative approach”, in Proceedings of IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA 2018), 20-22 April 2018, Chengdu, China.

[34] D. Huang, “Mobile cloud computing,” IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, vol. 6, no. 10, pp. 27–31, 2011.

[35] G. Md Whaiduzzaman et al, “A Study on Strategic Provision of Cloud Computing Services”, The Scientific World Journal, pp. 1-8, June 2014.

[36] S. K. Garg, S. Versteeg, R. Buyya, “A framework for ranking of cloud computing services”, Future Generation Computer Systems, vol. 29, issue: 4, pp.012–1023, 2013.

[37] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI: 10.1016/j.future.2016.11.031].

[38] Mohammad Haghighat et al, "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification," Expert Systems with Applications, vol. 11, no. 42, pp. 7905-7916, November 2015.

[39] Randeep Kaur, Supiya Kinger, "Analysis of Security Algorithms in Cloud Computing", International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 3, pp. 171-176, March 2014.

[40] Yogesh Kumar, Rajiv Munjal, Harsh Sharma, “Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures”,

IJCSMS International Journal of Computer Science and Management Studies, Vol. 11, Issue 03, October 2011.

[41] Gurpreet Singh, Supriya Kinger, “Integrating AES, DES, and 3-DES Encryption Algorithms for Enhanced Data Security”, International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July 2013.

[42] Abha Sachdev, Mohit Bhansali, “Enhancing Cloud Computing Security using AES Algorithm”, International Journal of Computer Applications, vol. 9, no. 67, pp. 19-23, April 2013.

[43] D. S. Abdul. Elminaam, H. M. Abdul Kader, M. M. Hadhoud, “Performance Evaluation of Symmetric Encryption Algorithms”, Communications of the IBIMA Volume 8, 2009.

[44] Uma Somani, “Implementing Digital Signature with RSA Encryption Algorithm to Enhance the Data Security of Cloud in Cloud Computing," 2010 1st International Conference on Parallel, Distributed and Grid Computing (PDGC-2010).

[45] J. Chen, Y. Chen, X. Du, C. Li, J. Lu, “Big data challenge: a data management perspective”, Springer, Frontiers of Computer Science, vol. 7, issue: 2, pp. 157-164, April 2013.

[46] C. Stergiou, K. E. Psannis, “Algorithms for Big Data in Advanced Communication Systems and Cloud Computing”, in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece. [DOI: 10.1109/CBI.2017.28]

[47] B. Marr, “Big Data: The 5 Vs Everyone Must Know”, LinkedIn article, 6 March 2014. Retrieved: 17/12/2018. Link: https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know

[48] S. Kaisler, F. Armour, J. A. Espinosa, W. Money, “Big Data: Issues and Challenges Moving Forward”, in Proceedings of 46th Hawaii International Conference on System Sciences, 7-10 January 2013, Wailea, Maui, HI, USA. [DOI: 10.1109/HICSS.2013.645]

[49] K. Bakshi, “Considerations for big data: Architecture and approach”, in Proceedings of 2012 IEEE Aerospace Conference, 3-10 March 2012, Big Sky, MT, USA. [DOI: 10.1109/AERO.2012.6187357]

[50] JASON, “Data Analysis Challenges”, The MITRE Corporation, Authorized to DOD and Contractors; Specific Authority, JSR-08-142, December 2008.

[51] C. Yang, M. Goodchild, Q. Huang, D. Nebert, R. Raskin, Y. Xu, M. Bambacus, D. Fay, “Spatial Cloud Computing: How Can the Geospatial Sciences Use and Help Shape Cloud Computing?”, International Journal of Digital Earth, vol. 4, issue: 4, pp. 305–329, June 2011. [DOI: 10.1080/17538947.2011.587547]

[52] J. A. Benediktsson, J. Chanussot, W. M. Moon, “Advances in Very-High-Resolution Remote Sensing”, in Proceedings of the IEEE, vol. 101 issue: 3, pp. 566–569, March 2013. [DOI: 10.1109/JPROC.2012.2237076]

[53] P. N. Edwards, “A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming”, 518. Cambridge, MA: MIT Press, 2010.

[54] J. L. Schnase, D. Q. Duffy, G. S. Tamkin, D. Nadeau, J. H. Thompson, C. M. Grieg, M. A. Mclnerney, W. P. Webster, “MERRA Analytic Services: Meeting the big Data Challenges of Climate Science Through Cloud-Enabled Climate Analytics-as-A-Service”, Computers, Environment and Urban Systems, vol. 61, Part B, pp. 198-211, January 2017. [DOI:10.1016/j.compenvurbsys.2013.12.003]

[55] L. Atzori, A. Iera, C. Morabito, "The Internet of Things: A survey", Computer Networks, vol. 54, issue: 15, pp. 2787–2805, October 2010. [DOI: 10.1016/j.comnet.2010.05.010]

[56] S. Lohr, “The age of big data”, New York Times, 11 February 2012. Retrieved: 10/01/2019. Link: https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html

[57] L. M. Camarinha-Matos, S. Tomic, P. Graça, (Eds.). 2013. “Technological Innovation for the Internet of Things”, in Proceedings of 4th IFIP WG 5.5/SOCOLNET, Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2013, Costa de Caparica, Portugal, April 15–17, 2013, Vol. 394, Springer.

[58] K. Michael, K. Miller, “Big Data: New Opportunities and New Challenges [Guest Editors” Introduction]”, IEEE, Computer, vol. 46, issue: 6, pp. 22–24, June 2013.

[59] R. Van den Dam, “Internet of Things: The Foundational Infrastructure for a Smarter Planet”. In Proceedings of Conference on Internet of Things and Smart Spaces International Conference on Next Generation Wired/Wireless Networking, Springer, ruSMART 2013, NEW2AN 2013: Internet of Things, Smart Spaces, and Next Generation Networking, pp. 1–12, vol. 8121, Berlin, Heidelberg, Germany, 2013.

[60] J. Belissent, “Getting Clever About Smart Cities: New Opportunities Require New Business Models”, For CIOs, 2 November 2010. Retrieved: 25/11/2018.

Likn1: http://193.40.244.77/iot/wp-content/uploads/2014/02/getting_clever_about_smart_cities_new_opportunities.pdf Link2: https://www.forrester.com/report/Getting+Clever+About+Smart+Cities+New+Opportunities+Require+New+Business+Models/-/E-RES56701

[61] H. Schaffers, N. Komninos, M. Pallot, B. Trousse, M. Nilsson, A. Oliveira, "Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation", Springer, Future Internet Assembly, FIA 2011, vol. 6656, pp. 431–446, 2011.

[62] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions", Elsevier, Future Generation Computer Systems, vol. 29, issue: 7, pp. 1645–1660, September 2013.

[63] Internet Live Stats, 1 July 2013. Retrieved 23/09/2018. Link: http://www.internetlivestats.com/internet-users/

[64] D. Boyd, K. Crawford, "Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon", Journal of Information, Communication & Society, vol. 15, issue: 5, pp. 662-679, May 2012.

[65] H. Chen, R. L. Chiang, V. C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact", MIS Quarterly: Management Information Systems, vol. 36, issue: 4, pp. 1165-1188, December 2012.

[66] V. Gopalkrishnan, D. Steier, H. Lewis, J. Guszcza, "Big data, big business: bridging the gap", in Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications, BigMine '12, pp. 7-11, 12 August 2012, Beijing, China. [DOI: 10.1145/2351316.2351318]

[67] F. Farber, S. K. Cha, J. Primsch, C. Bornhovd, S. Sigg, W. Lehner, "SAP HANA database: data management for modern business applications", ACM, SIGMOD Record, vol. 40, issue: 4, pp. 45-51, December 2011.

[68] A. B. M. Moniruzzaman, S. A. Hossain, "NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison", International Journal of Database Theory and Application, vol. 6, issue: 4, pp. 1-14, June 2013.

[69] C.-H. Hsu, K. D. Slagter, Y.-C. Chung, "Locality and loading aware virtual machine mapping techniques for optimizing communications in MapReduce applications", Elsevier, Future Generation Computer Systems, vol. 53, pp. 43-54, December 2015. [DOI: 10.1016/j.future.2015.04.006]

[70] L. Duan, W. N. Street, E. Xu, "Healthcare information systems: data mining methods in the creation of a clinical recommender system", Journal of Enterprise Information Systems, vol. 5, issue: 2, January 2011. [DOI: 10.1080/17517575.2010.541287]

[71] P. O'Donovan, K. Leahy, K. Bruton, D. T. J. O'Sullivan, "Big data in manufacturing: a systematic mapping study", Springer, Journal of Big Data, vol. 2, issue: 20, December 2015. [DOI: 10.1186/s40537-015-0028-x]

[72] H. Sequeira, P. Carreira, T. Goldschmidt, P. Vorst, "Energy Cloud: Real-Time Cloud-Native Energy Management System to Monitor and Analyze Energy Consumption in Multiple Industrial Sites", in Proceedings of 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing, 8-11 December 2014, London, UK. [DOI: 10.1109/UCC.2014.79]

[73] Z. Gui, M. Yu, C. Yang, S. Chen, J. Xia, Q. Huang, K. Liu, Z. Li, M. A. Hassan, B. Jin, "Developing Subdomain Allocation Algorithms Based on Spatial and Communicational Constraints to Accelerate Dust Storm Simulation", PLoS One, vol. 11, issu:4, e0152250, April 2016. [DOI: 10.1371/journal.pone.0152250]

[74] H. Kagermann, J. Helbig, A. Hellinger, W. Wahlster, "Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry", Final Report of the Industrie 4.0 Working Group, Forschungsunion, pp. 1-82, April 2013.

[75] N. Leavitt, Storage challenge: where will all that big data go? Computer 46(2013) 22–25.

[76] D. Talia, "Clouds for scalable big data analytics", Computer, vol. 46, pp. 98–101, 2013.

[77] M. D. Assuncao, R. N. Calheiros, S. Bianchi, M. A. Netto, R. Buyya, Big Data Computing and Clouds: Challenges, Solutions ,and Future Directions, arXiv preprint arXiv:1312.4722, (2013).

[78] N. Park, M. Kim, H.-C. Bang, "Symmetric Key-Based Authentication and the Session Key Agreement Scheme in IoT Environment", Springer, Computer Science and its Applications, vol. 330, pp. 379-384, 2015.

[79] C. Stergiou, A. P. Plageras, K. E. Psannis, B. B. Gupta, "Secure Machine Learning scenario from Big Data in Cloud Computing via Internet of Things network", Springer, Handbook of Computer Networks and Cyber Security: Principles and Paradigms, Multimedia Systems and Applications, 2019.

[80] K. Gai, M. Qiu, H. Zhao, “Privacy-Preserving Data Encryption Strategy for Big Data in Mobile Cloud Computing”, IEEE Transactions on Big Data, pp. 1-1, in Press, May 2018. [DOI: 10.1109/TBDATA.2017.2705807]

[81] H. Matallah, G. Belalem, K. Bouamrane, “Towards a New Model of Storage and Access to Data in Big Data and Cloud Computing”, International Journal of Ambient Computing and Intelligence (IJACI), vol. 8, issue: 4, pp. 31-45, October-December 2017. [DOI: 10.4018/IJACI.2017100103]

[82] S. K. Mishra, D. Puthal, B. Sahoo, S. K. Jena, M. S. Obaidat, “An adaptive task allocation technique for green cloud computing”, Journal of Supercomputing, vol. 74, issue: 1, pp. 370-385, January 2018. [DOI: 10.1007/s11227-017-2133-4]

[83] R. Chaudhary, G. S. Aujla, N. Kumar, J. J. P. C. Rodrigues, “Optimized Big Data Management across Multi-Cloud Data Centers: Software-Defined-Network-Based Analysis”, IEEE Communications Magazine, vol. 56, issue: 2, pp. 118-126, February 2018. [DOI: 10.1109/MCOM.2018.1700211]

[84] Y. Wen, J. Liu, W. Dou, X. Xu, B. Cao, J. Chen, “Scheduling workflows with privacy protection constraints for big data applications on cloud”, Elsevier, Future Generation Computer Systems, in Press, March 2018. [DOI: 10.1016/j.future.2018.03.028]

Chapter 6

Recent advances delivered in Mobile Cloud Computing’s Security and Management challenges

C. Stergiou, K. E. Psannis, “Recent advances delivered in Mobile Cloud Computing’s Security and Management challenges”, IGI Global, Modern Principles, Practices, and Algorithms for Cloud Security, 2019.

6.1 Introduction

Mobile cloud computing provides an opportunity to restrict the usage of huge hardware infrastructure and to provide access to data, applications, and computational power from every place and at any time with the use of a mobile device. Furthermore, MCC offers many possibilities but additionally creates several challenges and issues that need to be addressed as well. Through this work, the authors try to define the most important issues and challenges in the field of MCC technology by illustrating the most significant works related to MCC during recent years. Regarding the huge benefits offered by the MCC technology, this research tries to achieve a more safe and trusted environment for MCC users to operate the functions and transfer, edit, and manage data and applications, proposing a new method based on the existing AES encryption algorithm, which is, according to the study, the most relevant encryption algorithm to a cloud environment. Concluding, this research suggests a future plan to focus on finding new ways to achieve better integration of MCC with other technologies.

6.2 Related Work

For the purpose of this paper we study and analyze previous literature which has been studying two aspects, Security and Privacy Management in Mobile Cloud Computing [23-34] and Security and Privacy of Mobile Cloud Computing [35-44]. In addition to this, we also present some former works of our research group which have been made in field of Cloud Computing in general [3] [5] [6] [11] [45] [46]. All the papers presented with ascending form, from the older to the newest. The following paragraphs present the papers which contributed significantly in our study.

6.2.1. Security & Privacy Management in MCC

Initially, the papers that deal with the Security and Privacy issues of Management in MCC are illustrated [23-34]. As we can realize there are several works in this field. More particular, in [23] the authors propose an entity-centric approach for an IDM model in Cloud environment. The proposed approach based on two aspects: a) active bundles, and b) anonymous identification. The active bundles include a payload of Personally Identifiable Information, privacy policies and a virtual machine that enforces the policies and additionally the active bundles use a set of protection mechanisms in order to protect themselves. As regard the anonymous identification, they use it with the aim to mediate interactions between the entity and the Cloud services using entity's privacy policies. Moreover, the authors present the main characteristics of the approach which are: a) independent of third party, b) provides minimum information to the Service Provider, and c) provides ability to use identity data on untrusted hosts. Then, the [24] demonstrates the implementation of a mobile system that enables electronic healthcare data storage, update and retrieval using Cloud Computing. The proposed mobile application based in Google's Android OS and offers management of patient health records and medical images. This system was evaluated with the use of Amazon’s S3 cloud service. Finally, the authors summarize the details of the implementation and then present initial results of the system in practice. Moreover, the authors of [25] survey the MCC technology, which could help the general readers to have an overview of the MCC including the definition, the architecture, and the applications. Also, the [25] presents the issues, the existing solutions, and the recent approaches of the MCC technology. At the end, the authors discuss a number of future research directions of the MCC. Through the [26] the authors propose a multi-faceted Trust Management system architecture for a cloud computing marketplace, with the aim to support the customers in reliably identifying trustworthy cloud providers. The proposed system offers means to identify the trustworthy cloud providers in term of different attributes that assessed by multiple sources and roots of trust information. Furthermore, the [27] presents a sort survey of MCC evolution and additionally explains how Cloud Computing and Mobile Devices could be combined with good terms for future opportunities, implications and legal issues for developing countries. In another research, the authors of [28] try to review the existing Distributed Application Processing Frameworks, also known as DAPFs, for SMDs in MCC domain. The main objective of [28] is to highlight issues and challenges to existing DAPFs in developing, implementing, and executing computational intensive mobile applications within MCC domain. Thus, through this work the authors propose a thematic taxonomy of the current DAPFs, and then they review current offloading frameworks by using thematic taxonomy, and analyze the implications and critical aspects of current offloading frameworks. Finally, the [28] puts forward open research issues in distributed application processing for MCC that remain to be addressed. Also, the [29] proposes a trust management approach by making an analysis of user behavioral patterns for a reliable Mobile Cloud Computing. So, the authors suggest a method in order to quantify a one-dimensional trusting relation count on the analysis of telephone call data from Mobile Cloud Environment. Subsequently, it is enhanced trustworthiness of data production, management, and overall application. In [30] it was presented a state-of-the-art survey of vehicular Cloud Computing. More detailed, the authors present a taxonomy for vehicular Cloud in which special attention has been devoted to the extensive applications, Cloud formations, key management, inter-Cloud communication systems, and broad aspects of privacy and security problems. Additionally, in [30] the authors illustrates the design of architecture for Vehicular Cloud Computing, itemize the properties required in vehicular Cloud which support the proposed model. In order to achieve their goal, authors compare the proposed mechanism with normal Cloud Computing and then discuss about open research issues and the future directions. Additionally, in [31] the authors discuss the limitations of the state-of-the-art Cloud Identity Managements with respect to mobile clients. In particular, the authors demonstrate that the current IDMs are vulnerable to three attacks. As a result of their research, the authors propose and validate a new IDM architecture dubbed Consolidated IDM that countermeasures these attacks. Through their experimental results the authors illustrates that CIDM offers its clients with better security guarantees and that it has less energy and communication overhead compared to the current IDM systems. Furthermore, the [32] offers a detailed survey of the security issues that arise due to the very nature of Cloud Computing. Furthermore, the [32] presents a number of recent solutions offered in the literature with the aim to counter the security problems. In addition, there is a brief view of the highlighted security vulnerabilities in the Mobile Cloud Computing. Then, in [33] there is a discussion about the evolution of the computing as regarding the historical perspective, with focus on advances which primarily led to the evolution of the Cloud Computing. Additionally, there is a survey of some of the critical components that are vital in order to make the Cloud Computing paradigm feasible. The authors by addressing a number of particular legal and philosophical problems try to conclude with a hard look at all the successful Cloud Computing vendors. Finally, the authors through the [34] try to present the major security and privacy issues and challenges in the field which have grown much interest among the academia and research community. Also, they try to illustrate reports of recent works of literature. Moreover, they present a comparative analysis of the literature works based on different security and privacy requirements, and concluding they present the open issues in the field.

6.2.2. Security & Privacy of MCC

Subsequently, the papers that deal with the Security and Privacy issues of the Mobile Cloud Computing are illustrated [35-44]. Starting from the oldest, a survey of the various security issues that pose a threat to the Cloud presented in [35]. The survey that illustrated in [35] could be more specific to different security problems which have emanated due to the nature of the service delivery models of a Cloud Computing system. Continuously, the authors of [36] propose a Security Service Admission Model based on Semi-Markov Decision Process in order to model the system reward for the Cloud provider. Initially, through the [36] try to define the system states by a tuple represented by the numbers of Cloud users and the associated to them security service categories, and the current event type. Then, the authors derive the system steady-state probability and service request blocking probability with the use of their proposed model. Then, the [37] proposes a trust model for Cloud architecture which uses mobile agent as security agents to acquire useful information from the virtual machine that the user and the service provider could utilize with the aim to keep track of privacy of their data and virtual machines. In the proposed model the security agents can dynamically move through the network, replicate themselves according to the requirement, and perform the assigned tasks like accounting and monitoring of virtual machines. Also, in [38] the authors propose a framework with the aim to secure the data transmitted between the components of the same Mobile Cloud Application, and with the aim to ensure the integrity of the applications at the installation on the mobile device and when being updated. Furthermore, the proposed framework of [38] allows applying different security properties to various kinds of data and not the same properties to all the data processed by the application. In addition to this, the proposed approach takes into account the user's preferences and the mobile device performances. Moreover, the [39] discusses and identifies the main vulnerabilities in Cloud Computing systems, and the most important threats that found in the literature related to Cloud Computing technology and its environment, as well as to identify and relate vulnerabilities and threats with possible solutions. In [40] the authors with the aim to facilitate the emerging domain of MCC security and privacy, in brief review the advantages and system model of MCC, and pay attention to the security and privacy in the MCC. At the end, the authors provide the current security and privacy approaches, by deeply analyzing the security and privacy problems from the aspects of mobile terminal, mobile network and Cloud. Additionally, the [41] presents a comprehensive literature review of MCC and the security and privacy issues that MCC faced. Also, the authors of [41] present a complete understanding analysis of MCC, in where they explain its architecture, advantages and applications. At the end, the authors conclude that their research is significant useful for the mobile service providers, and thus they can improve the security technologies and mechanisms used for Cloud security in order to minimize the user's security concerns. Then, in [42] illustrated particular efforts that have been devoted in research organizations and academia in order to build secure Mobile Cloud Computing environments and infrastructures. Thus, in the spite of the efforts, there are a number of loopholes and challenges that still exist in the security policies of MCC. The authors through the literature review conclude in three things that discussed in [42]. Firstly, they highlight the current state of the art work which proposed to secure MCC infrastructures. Secondly, they identify the potential problems. Finally, they provide a taxonomy of the state of the art. Furthermore, the authors of [43] propose a novel technique that called "match-then-decrypt", in which a matching phase is also presented before the decryption phase. The proposed technique operates by computing special components in ciphertexts, which are used in order to perform the test that if the attribute private key matches the hidden access policy in ciphertexts without decryption. Moreover, in [43] there is a proposal of a basic anonymous ABE construction, and then obtain a security-enhanced extension based on strongly existentially unforgeable one-time signatures. More specifically, the authors conclude that the formal security analysis and performance comparisons indicate that their proposed solutions simultaneously ensure attribute privacy and improve decryption efficiency for outsourced data storage in MCC. Finally, the [44] initially identify that the scheme that proposed in former work of Tsai and Lo, which was a privacy aware authentication scheme for distributed MCC services, fails to achieve mutual authentication. The authors of [44] conclude to this regarding it is vulnerable to the service provider impersonation attack. In addition to this, the authors state that the former scheme also suffers from some minor design flaws, including the problem of biometric measure, wrong password, and fingerprint login, no user revocation facility when the smart card is lost or stolen. At the end, they offer some a number of suggestions in order to avoid the aforementioned design flaws in future design of authentication schemes.

6.2.3. Research group's previous works in CC & MCC

At this point there will be present a number of former works which deal with problems and solutions in the field of Cloud Computing in general [3] [5] [6] [11] [45] [46]. More particular, some of them deal with problems and solutions in the field of Mobile Cloud Computing. Starting again from the oldest, in [5] the authors try to combine the MCC and IoT with the Big Data with the aim to examine the common characteristics and in addition to discover which of MCC and IoT benefits improve the operation of Big Data applications. Also, the authors of [5] present the contribution of MCC and IoT individually to Big Data. Moreover, the authors of [45] present a survey of Internet of Things and Cloud Computing focusing on the security problems of both of them. More particular, the authors try to combine these technologies aiming to examine the common features, and also aiming to discover the benefits of their integration. At the end, there is a presentation of the contribution of Cloud Computing to the IoT. So, the [45] illustrates how the Cloud Computing improves the functionality of Internet of Things, and additionally, surveys the security challenges of the integration of Cloud Computing and Internet of Things. Continuously, the [6] present a survey of Big Data and Cloud Computing, illustrating their basic characteristics, and focusing on the security and privacy problems of both of them. Regarding this, the authors try to combine the functionality of Big Data and Cloud Computing aiming to examine the frequent characteristics, and in addition to this to discover the benefits which are related in security problems of their integration. Furthermore, the authors of [14] survey Cloud Computing and Big Data technology, and their major characteristics, focusing on the security and privacy problems of both of them. Particularly, the authors combine the functionality of two technologies aiming to examine the common characteristics, and additionally to discover the benefits related in security issues of their integration. Then, there is a presentation of a novel method of an algorithm that can be used for the purpose of improving Cloud Computing's security through the use of algorithms that can offer more privacy in the data related to Big Data. At the end, there additionally a survey about the challenges of the integration of Cloud Computing and Big Data related to their security level. Also, in [3] the authors in order to achieve a type of network that will offer more intelligent media-data transfer new technologies were studied. Thus, the authors initially studied the use of various open source tools of Cloud Computing analyzers and simulators. So, the authors after they measure the simulated network performance with CloudSim simulator, they use the Cooja emulator of the Contiki OS aiming to confirm and access more metrics and options. In particular, in [3] there is an implementation of a network topology from a small section of the script of CloudSim with Cooja, so that the authors could test a single network segment. The results that have been produced of the experimental procedure illustrate that there are not duplicated packets received during the whole procedure. Finally, in [46] the authors propose a novel system for Cloud Computing integrated with Internet of Things as a base scenario for Big Data. Moreover, the authors try to establish an architecture relaying on the security of the network with the aim to eliminate the security issues.

The solution proposed in [46] installs a security "*wall*" between the Cloud server and the outer Internet, aiming to eliminate the privacy and security problems. As regard the main goal of [46] a sort survey of IoT and Cloud Computing also presented, focusing on the security issues of both of them. Additionally, the authors state that through their study conclude that Cloud Computing could offer a more "*green*" and efficient "*fog*" environment for sustainable computing scenarios.

6.2.4. Literature Comparative Analysis

Taking into account the Related Research Review Section we realized that the study of MCC's Security and Privacy issues become more popular in the research and academic community over the recent years. We have come to the above conclusion counting on our study of several works in the field of Security and Privacy of Mobile Cloud Computing. The main balk of these works presented in Related Research Review Section. Consequently, there is a need for further research in this area as the growing numbers studied indicate.

As we can observe, the last four years the interest of the researchers has increased considerably compared to the previous decade. Figure 6.1 illustrates the growth of studies in Security and Privacy Management in Mobile Cloud Computing through the years. Equally important Figure 2 represents the growth of studies in Security and Privacy of Mobile Cloud Computing through the years.

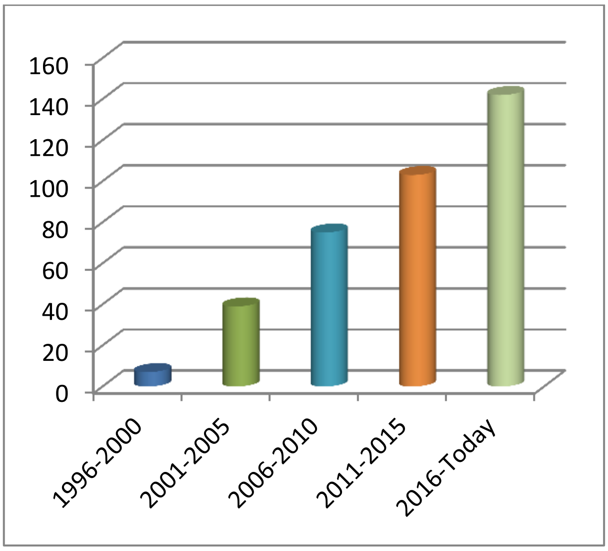


Figure 6.1: Growth of works made in Security and Privacy Management in Mobile Cloud Computing over the years

Year	Author	Challenge/Issue
2009-08	J. W. Rittinghouse & J. F. Ransome **[33]**	• Cloud reliability to users • Users desired outcomes of the Cloud • Levels of trust in Cloud environment
2010-09	C. Doukas et al **[24]**	• Sharing and management of medical information resources through mobile healthcare systems
2010-11	P. Angin et al **[23]**	• Entities authentication to service providers • Entities multiple accounts associated with multiple service providers
2011-01	S. Subashini & V. Kavita **[35]**	• Cloud environment safety • New Cloud model targeting to improve the existing one
2011-04	H. Liang et al **[36]**	• Increased number of Critical and Normal Security service users • Allocate Cloud resource aiming to maximize the system rewards with the considerations of the Cloud resource consumption and incomes generated from Cloud users.
2011-10	H. T. Dinh et al **[25]**	• Low bandwidth of Mobile Cloud Computing • Availability of Mobile Cloud Computing • Heterogeneity of Mobile Cloud Computing • Computing offloading of Mobile Cloud Computing • Security of Mobile Cloud Computing • Enhancing the efficiency of data access of Mobile Cloud Computing • Context-aware mobile Cloud services of Mobile Cloud Computing
2011-11	S. M. Habib et al **[26]**	• Not consistent descriptions in Service Level Agreements among Cloud providers • Uncertain reliably identifying trustworthy Cloud providers for the customers
2011-12	P. S. Hada et al **[37]**	• Major need of bringing reliability, transparency and security in Cloud model for client satisfaction
2012-10	M. R. Prasad et al **[27]**	• Cloud Computing challenges: Performance, Security & Privacy, Control, Bandwidth Costs, Reliability • Mobile Cloud Computing legal issues
2012-11	M. Shiraz et al **[28]**	• Users expectation to run computational intensive applications on Smart Mobile Devices in the same way as powerful stationary computers • Establishment of distributed application processing platform at runtime which requires additional computing resources on Smart Mobile Devices
2013-01	D. Popa et al **[38]**	• Use of MCC increases security risks and privacy invasion • MCC application model not well-defined
2013-06	A. Shahzad & M. Hussain **[41]**	• Security and privacy risks faced be MCC uses • Architecture and Cloud service delivery models issues • Mobile Cloud infrastructure issues • Mobile Cloud communication channel issues
2013-07	H. Suo et al **[40]**	• Security and privacy issues and challenges in MCC
2013-07	A. N. Khan et al **[42]**	• Security threats have become a hurdle in the rapid adaptability of the MCC paradigm
2013-12	M. Kim & S. O. Park **[29]**	• MCC architecture, design and implementation need to be improved due to limited computing capability and storage issues • Trust management approach for reliable MCC
2013-12	K. Hashizume et al **[39]**	• Risk of outsourcing data to third party Cloud providers • Cloud Computing inherits many technologies security issues
2014-02	M. Ali et al **[32]**	• Services provided by third party Cloud providers entail additional security threats • Migration of user's assets outside the administrative control in the Cloud environment escalates the security concerns.
2014-03	I. Khalil et al **[31]**	• Mobile devices are easy to be compromised • Mobile users store Personal Identifiable Information in unprotected text files, cookies and applications • Limitations of state-of-the-art Cloud Identity Management Systems with respect to mobile clients
2014-04	Md. Whaiduzzaman et al **[30]**	• Traffic management and road safety by instantly using vehicular resources • Vehicular Cloud Computing security challenges: Authentication, Secure location & localization, Securing vehicular communication, Vehicular public key infrastructure, Data security, Network heterogeneity, Access control
2017-02	Y. Zhang et al **[43]**	• Each decryption usually requires many pairings and the computation overhead grows with the complexity of the access formula • Existing schemes suffer a serve efficiency drawback and not suitable for MCC where users may be resource-constrained
2017-04	M. B. Mollah et	• Data security challenges

	all **[34]**	• Partitioning and offloading security challenges • Virtualization security challenges • Mobile Cloud applications security challenges • Mobile devices security challenges • Privacy challenges
2018-06	Q. Jiang et al **[44]**	• A previous proposed scheme of Tsai & Lo fails to achieve mutual authentication and withstands all major security threats

Table 6.1: Related research work's challenges and issues of the Mobile Cloud Computing technology

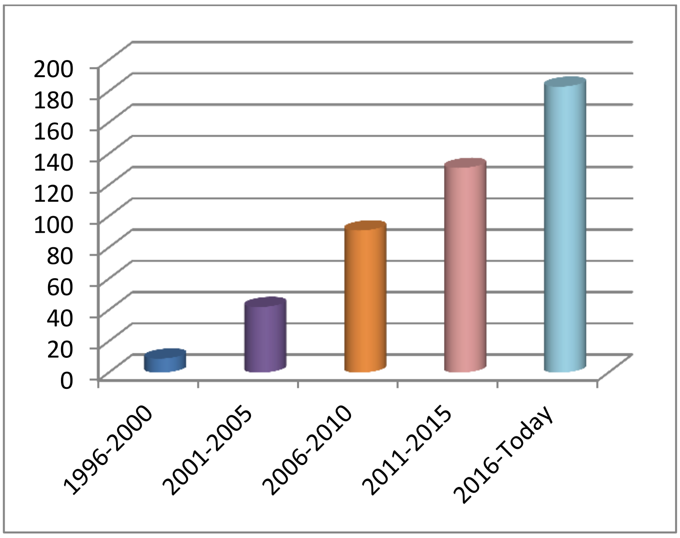


Figure 6.2: Growth of works made in Security and Privacy of Mobile Cloud Computing over the years

The study of previous works motivates us to survey the Security and Privacy challenges and issues of the Mobile Cloud Computing technology. Thus, count on the major works related to Security and Privacy challenges and issues we have tried to figure out them in Table 6.1.

Table 6.1 lists the major challenges and issues which we have distinguished from the related works. Through Table 6.1 we can figure out which are the major issues and challenges of the Mobile Cloud Computing that have been addressed by the literature.

6.3 Research Outcomes & Proposed Solutions

Literature work	*Challenges*								
	Privacy	*Security*	*Trusted environment*	*Bandwidth*	*Environment limitations*	*Management*	*Reliability*	*User authentication*	*Efficiency*
Rittinghouse & Ransome [33]			X			X	X	X	
Doukas et al. [24]			X			X			
Angin et al. [23]	X	X				X		X	
Subashini & Kavitha [35]			X		X	X			
Liang et al. [36]		X	X			X	X	X	
Dinh et al. [25]		X	X	X			X		X
Habib et al. [26]		X	X				X	X	
Hada et al. [37]		X	X				X		
Prasad et al. [27]	X	X		X		X	X	X	
Shiraz et al. [28]					X	X			X
Popa et al. [38]	X	X			X	X			
Shahzad & Hussain [41]	X	X	X		X		X		
Suo et al. [40]	X	X	X						
Khan et al. [42]		X			X				
Kim & Park [29]			X	X	X		X		X
Hashizume et al. [39]		X	X				X	X	
Ali et al. [32]		X	X		X	X	X		
Khalil et al. [31]			X		X	X	X	X	
Whaiduzzaman et al. [30]		X				X	X	X	X
Zhang et al. [43]			X	X	X				X
Mollah et al. [34]	X	X		X					
Jiang et al. [44]		X	X		X			X	

Table 6.2: Literature challenges in the field of Mobile Cloud Computing

Table 6.2 presents the challenges that have been addresses by the literature work which we have studied. As we could observe most of the works that we have studied focus on the "*Trusted environment*" which in our opinion is the major issue of the Mobile Cloud Computing, and as a result the "*Security*" issue is most popular in the literature work. Additionally, concerning the rest of the challenges we come to the conclusion that the "*Reliability*" and the "*Management*" are also basic issues that need to be addressed. More detailed, through the number of 22 works that we have stood out the statistic results are the following: Privacy 6 of 22, Security 15 of 22, Trusted environment 15 of 22, Bandwidth 5 of 22, Environment limitations 10 of 22, Management 11 of 22, Reliability 12 of 22, User authentication 9 of 22, Efficiency 5 of 22. Equally important, the less mentioned issues are the "*Bandwidth*" and the "*Efficiency*" which are really vital for the functionality of the Mobile Cloud Computing technology. Regarding this, we could realize that more researches need to be done in the field in order to find better solutions aiming to improve Bandwidth and Efficiency of MCC.

Thus, based on previous works [3] [6] [7] [11] [13] [14] [22] [45] [46] the optimal solution in order to achieve a reliable and trusted environment with a more "*safe*" authentication and encryption for the users the MCC system have to adapt the AES encryption algorithm.

The AES algorithm has variable key length of 128, 192, or 256 bits, with default 256 bits. It encrypts data blocks of 128 bits in 10, 12 and 14 round depending on the key size. AES encryption is fast and flexible algorithm, and it can be implemented on various platforms especially in small devices, such as mobile devices. Also, AES has been carefully tested for many security applications [22]. As a result, the AES algorithm provides the ability to have speed key setup time a good key agility. So, if we use this algorithm we could have a trusted relay method with an encryption of a speed key setup. Therefore, instead the trust relay use we can seize also there no serious weak keys in AES and so we could have a beneficial security use of the encryption in the integrated new model. Similarly, we can take advantage the less memory which AES needs for implementation that makes it for restricted-space environments. Thus, we can seize the transmit power that AES offers and as a result we can have a better and more trusted transmission in a MCC environment [14] [22] [46].

Count our study and the reliability of AES encryption algorithm for a MCC environment we suggest the following part of pseudocode based on the AES encryption algorithm.

Algorithm 1

```
input -> byte[]
byte[] + R.Key -> state[]
for 6 to 66
  W[i-1] -> T
  if i mod 6 = 0
                rotate T + 6
       W[i-6] / T -> W[i]
       R.Key+1
       i+1 -> i
Row +1 -> Row
state[] -> output[]
```

Table 6.3: Suggested pseudocode based on the AES encryption algorithm

With this proposed method we can extend the advances of MCC, in particular when it integrates with other technologies like IoT and Big Data, by developing a highly innovative and scalable service platform to enable secure and privacy services. Through tour research we can propose the algorithm 1 which extends the security advances of MCC environment, especially when integrates with other technologies. As a proposal of this work could be this part of pseudocode algorithm which uses the original key consists of 128 bits/16 bytes which are represented as a 6x6 matrix.

6.3.1. Experimental Results

Compared our proposed method with the existing one we come to some measurements that have been through time and showing the benefits of our proposed method.

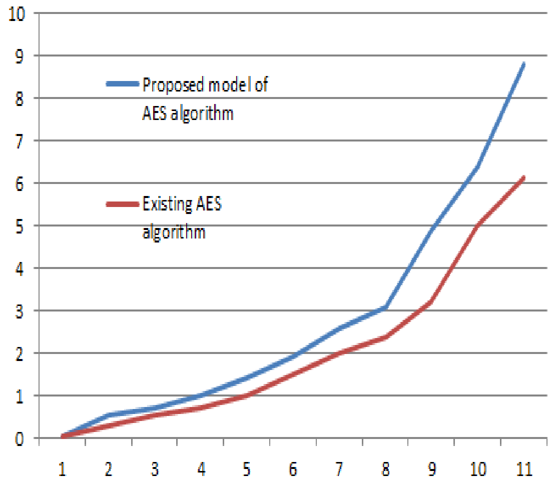


Figure 6.3: Security level of encryption algorithms of measurement used for the study of AES model algorithm

As we can observe by Figure 6.3 the more often is the combined use of the algorithms, the higher level of security of the data usage we get every time. The upper line represents our proposed model of AES algorithm and the other (down line) represents the existing AES algorithm. Based on Figure 6.3 we can also figure out that our proposed method is over from the existing model regarding the higher security level of encryption that it can achieve through the time.

6.4 Chapter Summary

The MCC technology provides a number of possibilities, but additionally places several challenges and issues that need to be addressed as well. Mobile Cloud Computing refers to an infrastructure where data, applications and information could be processed through a mobile device, but simultaneously outside of the mobile device. The main objective of the use of MCC is to decrease the use of stronger hardware and to have the access to data and applications, and in many times to more computational power, from every place and in any time, through a mobile device.

With our study, in regard on the huge benefits of the Mobile Cloud Computing technology, we try to achieve a more safe and trusted environment for the MCC users in order to operate the functions, and transfer, edit and manage data and applications. This could be achieved proving a novel method count on the AES encryption algorithm, which is, according to our study, the most relevant encryption algorithm to a Cloud environment.

Furthermore, we try to define the most important issues and challenges in the field of Mobile Cloud Computing technology by presenting a number of the most significant works related to MCC through the last eight years.

As a future work, we could focus to find novel ways to achieve a better integration MCC with other technologies, focusing on security algorithms and all the challenges that the technologies faced on security level. Regarding the rapid development of Cloud technology the security issues of Mobile Cloud Computing must be solved or reduced to a minimum in order to have a better and safer model. The security challenges and issues that surveyed in this work could be the sector for further research as a case study, with the goal of minimizing them.

6.5 Chapter References

[3] C. Stergiou, K. E. Psannis, A. P. Plageras, Y. Ishibashi, B.-G. Kim, "Algorithms for efficient digital media transmission over IoT and cloud networking", Journal of Multimedia Information System, vol. 5, no. 1, pp. 27-34, March 2018. [DOI: 10.9717/JMIS.2018.5.1.27]

[5] C. Stergiou, K. E. Psannis, "Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey", Wiley, International Journal of Network Management, pp. 1-12, May 2016. [DOI:10.1002/nem.1930]

[6] C. Stergiou, K. E. Psannis, "Algorithms for Big Data in Advanced Communication Systems and Cloud Computing", in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece. [DOI: 10.1109/CBI.2017.28]

[7] C. Stergiou, K. E. Psannis, B. B. Gupta, "Advanced Media-based Smart Big Data on Intelligent Cloud Systems", IEEE Transaction on Sustainable Computing, in Press, 2018.

[11] C. Stergiou, K. E. Psannis, A. P. Plageras, T. Xifilidis, B. B. Gupta, "Security and Privacy of Big Data for Social Networking Services in Cloud", in Proceedings of IEEE conference on Computer Communications (IEEE INFOCOM 2018), 15-20 April 2018, Honolulu, HI, USA.

[13] C. Stergiou, K. E. Psannis, B. Gupta, Y. Ishibashi, "Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT", Elsevier, Sustainable Computing, Informatics and Systems, In Press, June 2018.

[14] C. Stergiou, K. E. Psannis, "Efficient and Secure Big Data delivery in Cloud Computing", Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017. [DOI: 10.1007/s11042-017-4590-4]

[22] G. Singh, S. Kinger, "Integrating AES, DES, and 3-DES Encryption Algorithms for Enhanced Data Security", International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July 2013.

[23] P. Angin, B. Bhargava, R. Ranchal, N. Singh, M. Linderman, "An Entity-centric Approach for Privacy and Identity Management in Cloud Computing", in Proceedings of 29th IEEE International Symposium on Reliable Distributed Systems, 31 October-3 November 2010, New Delhi, India. [DOI: 10.1109/SRDS.2010.28]

[24] C. Doukas, T. Pliakas, I. Maglogiannis, "Mobile Healthcare Information Management utilizing Cloud Computing and Android OS", in Proceedings of 32nd Annual International Conference of the IEEE EMBS 2010, 31 August-4 September 2010, Buenos Aires, Argentina.

[25] H. T. Dinh, C. Lee, D. Niyato, P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches", Wireless Communications and Mobile Computing, vol. 13, pp. 1587-1611, October 2011. [DOI: 10.1002/wcm.1203]

[26] S. M. Habib, S. Ries, M. Muhlhauser, "Towards a Trust Management System for Cloud Computing", in Proceedings of IEEE International Joint Conference TrustCom-11/IEEE ICESS-11/FCST-11, 16-18 November 2011, Changsha, China.

[27] M. R. Prasad, J. Gyani, P.R. K. Murti, "Mobile Cloud Computing: Implications and Challenges", Journal of Information Engineering and Applications, vol. 2, no. 7, pp. 7-15, October 2012.

[28] M. Shiraz, A. Gani, R. H. Khokhar, R. Buyya, "A Review on Distributed Application Processing Frameworks in Smart Mobile Devices for Mobile Cloud Computing", IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1294-1313, November 2012.

[29] M. Kim, S. O. Park, "Trust management on user behavioral patterns for a mobile cloud computing", Springer, Cluster Computing, vol. 16, issue 4, pp. 725-731, December 2013. [DOI: 10.1007/s10586-013-0248-9]

[30] Md. Whaiduzzaman, M. Sookhak, A. Gani, R. Buyya, "A survey on vehicular cloud computing", Elsevier, Journal of Network and Computer Applications, vol. 40, pp. 325-344, April 2014. [DOI: 10.1016/j.jnca.2013.08.004]

[31] I. Khalil, A. Khreishah, M. Azeem, "Consolidated Identity Management System for secure mobile cloud computing", Elsevier, Computer Networks, vol. 99, pp. 99-110, March 2014. [DOI: 10.1016/j.comnet.2014.03.015]

[32] M. Ali, S. U. Khan, A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges", Elsevier, Information Sciences, vol. 305, pp. 357-383, February 2015. [DOI: 10.1016/j.ins.2015.01.025]

[33] J. W. Rittinghouse, J. F. Ransome, "Cloud Computing: Implementation, Management, and Security", CRC Press, 340 Pages - 127 B/W Illustrations, 17 August 2009. ISBN 9781439806807 - CAT# K10347

[34] M. B. Mollah, Md. A. K. Azad, A. Vasliakos, "Security and privacy challenges in mobile cloud computing: Survey and way ahead", Elsevier, Journal of Network and Computer Applications, vol. 84, pp. 38-54, April 2017.

[35] S. Subashini, V. Kavitha, "A survey on security issues in service delivery models of cloud computing", Elsevier, Journal of Network and Computer Applications, vol. 34, issue: 1, pp. 1-11, January 2011. [DOI: 10.1016/j.jnca.2010.07.006]

[36] H. Liang, D. Huang, L. X. Cai, X. S. Shen, D. Peng, "Resource Allocation for Security Services in Mobile Cloud Computing", in Proceedings of IEEE Conference on Computer Communications Workshops INFOCOM WKSHPS 2011, 10-15 April 2011, Shanghai, China.

[37] P. S. Hada, R. Singh, M. M. Meghwal, "Security Agents: A Mobile Agent based Trust Model for Cloud Computing", International Journal of Computer Applications, vol. 36, no. 12, pp. 12-15, December 2011. [DOI: 10.5120/4547-6435]

[38] D. Popa, M. Cremene, M. Borda, K. Boudaoud, "A Security Framework for Mobile Cloud Applications", in Proceedings of 11th RoEduNet International Conference, 17-19 January 2013, Sinaia, Romania.

[39] K. Hashizume, D. G. Rosado, E. Fernandez-Medina, E. B. Fernandez, "An analysis of security issues for cloud computing", Springer, Journal of Internet Services and Applications, vol. 4, no. 5, pp. 1-13, December 2013. [DOI: 10.1186/1869-0238-4-5]

[40] H. Suo, Z. Liu, J. Wan, K. Zhou, "Security and Privacy in Mobile Cloud Computing", in Proceedings of 9th International Conference on Wireless Communications and Mobile Computing (IWCMC 2013), 1-5 July 2013, Sardinia, Italy.

[41] A. Shahzad, M. Hussain, "Security Issues and Challenges of Mobile Cloud Computing", International Journal of Grid and Distributed Computing, vol. 6, no. 6, pp. 37-50, 2013. [DOI: 10.14257/ijgdc.2013.6.6.04]

[42] A. N. Khan, M. I. Mat Kiah, S. U. Khan, S. A. Madani, “Towards secure mobile cloud computing: A survey”, Elsevier, Future Generation Computer Systems, vol. 29, issue: 5, pp. 1278–1299, July 2013.

[43] Y. Zhang, X. Chen, D. S. Wong, H. Li, I. You, “Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing”, Elsevier, Information Sciences, vol. 379, pp. 42-61, Februoary 2017. [DOI: 10.1016/j.ins.2016.04.015]

[44] Q. Jiang, J. Ma, F. Wei, “On the Security of a Privacy-Aware Authentication Scheme for Distributed Mobile Cloud Computing Services”, IEEE Systems Journal, vol. 12, issue: 2, pp. 2039-2042, June 2018. [DOI: 10.1109/JSYST.2016.2574719]

[45] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI: 10.1016/j.future.2016.11.031]

[46] C. Stergiou, K. E. Psannis, B. Gupta, Y. Ishibashi, “Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT”, Elsevier, Sustainable Computing, Informatics and Systems, In Press, June 2018. [DOI: 10.1016/j.suscom.2018.06.003]

Chapter 7

Secure Machine Learning scenario from Big Data in Cloud Computing via Internet of Things network

C. L. Stergiou, A. P. Plageras, K. E. Psannis, B. B. Gupta, "Secure Machine Learning scenario from Big Data in Cloud Computing via Internet of Things network", Springer, Handbook of Computer Networks and Cyber Security: Principles and Paradigms, Multimedia Systems and Applications, 2019.

Reprinted/adapted by permission from **C. L. Stergiou**: **Springer Handbook of Computer Networks and Cyber Security** by **B. B. Gupta, G. Martinez Perez, D. P. Agrawal, D. Gupta** chapter **Secure Machine Learning Scenario from Big Data in Cloud Computing via Internet of Things Network** by **C. L. Stergiou, A. P. Plageras, K. E. Psannis, B. B. Gupta 2019**

7.1 Introduction

Cloud Computing (CC) technology refers to an infrastructure in which both data storage and data processing takes place outside the mobile device. Furthermore, another new and fast-growing technology called the Internet of Things (IoT) raises in the sector of networks and telec.ommunications with specifical concern in the "modern" area of wireless telecommunications systems. Regarding our recent research, the main goal of the interaction and cooperation between things and objects sent through the wireless networks. It is to fulfill the objective set to them as a combined entity, to achieve a better environment for the use of Big Data (BD). In addition, count on the technology of wireless networks, both CC and IoT could be developed rapidly and together. In this paper, I survey IoT and Cloud Computing technologies with a focus on security problems that both technologies faced. Particularly, these two aforementioned technologies (i.e Cloud Computing and IoT) have been compared, aiming to the familiar characteristics, and examined and discover the benefits of their integration focusing to secure the use and transmission of Big Data. Concluding, a contribution of CC and IoT technologies have been presented, and how the CC technology improves the operation of IoT as base technologies for Big Data systems.

7.2 Related Work

To come through the proposed scenario various related works that discuss the combination of the three afore-mentioned technologies (Big Data, Cloud Computing and Internet of Things) have been studied. This section illustrates related work similar to this research. The main tumor of the related research studies is mainly related to previous work of our research team.

To start with, in [11] the authors aim in the interaction and the conjunction of Mobile Cloud Computing (MCC) and IoT through the integration of these technologies with the Big Data. This scenario, based on similar characteristics of MCC and IoT, and which of the benefits of these technologies could improve the use of BD applications. Also, in [11] an illustration has been presented of how the MCC and the IoT contribute to the BD technology, individually.

A region based research [2] presents a survey research of IoT and CC focusing on the issues based on data privacy of both technologies. Particularly, the authors of [2] try to combine these technologies with the purpose to find and examine the familiar characteristics and then discover the profits of their integration. Additionally, the authors illustrate the contribution of CC in the field of IoT, and through this it can be proved how the CC technology improves the operation of IoT.

In [7], the authors survey BD and CC technologies and their major features, focusing on security and data privacy issues. Particularly, a conjunction of the functionality of those two technologies has been done with the aim to consider the frequent characteristics, and in addition to this, to discover the profits which deal with security problems of their integration. Thus, a novel method of an algorithm has been presented in [7], which could be used for the purpose of upgrading the CC's security through the use of algorithms that can provide privacy of the large amounts of data.

Another research [8] focuses on a proposal of system integration between IoT and Video Surveillance (VS) technology, with the goal to indulge the requirements of the future needs of VS, and to accomplish a better use of it. The VS data that have been transmitted through the network could be characterized as large-scale data, and thus as BD. The basic outcome of the specific research [8] is an innovative topology paradigm which could offer a better use of IoT technology in VS, and vice-versa.

In [24] initially, it has been presented an analytical study of IoT, CC and BD to resolve various issues that face the health sector in regard to these technologies. In the proposed scenario there is a collection of e-health data by sensor devices and actuators which has been transferred through an established network to a cloud server. These data could be processed in the cloud server in or-der to be analyzed, and by this analysis there would be born what we call "*data mining*". Moreover, there is a research [24] that deals with security of medical data which constitute sensitive personal data and must be protected.

Moreover, in [3] the authors initially present a survey of the technologies IoT, BD, CC and Monitoring with the aim to discover their common operations and to combine their functionality, in order to achieve beneficial scenarios of their use. The main objective of [3] is to propose a novel system which operates in IoT environment, within there will be collected and managed sensors' data. Additionally, the authors state that their proposed system will be energy efficient and it would be used in a "*Green Smart Building*".

In [12] the authors try to achieve and propose a type of network that will provide more intelligent media-data transfer. Thus, through the study of the use of various open source tools, the authors found the suitable for their experiments tool with the aim to measure the performance of their proposed model of network. At the end, the authors proposed the network topology that they have implemented from a small section of the script of CloudSim simulator with Cooja, so that they could test a single network segment.

The [25] surveys Social Networking (SNg), BD and CC, focusing on their main features, by concentrating on the security problems of those technologies. In particular, the authors aim to combine the functionality of BD and SNg in CC environment, so that they could analyze the common characteristics and ascertain the advantages of their integration related to security issues. The main outcome of [25] is the presentation of a novel system-framework-network in Cloud environment through which users of various Social Networks (SNs) will be able to exchange data and information, and primarily large-scale data.

To summarize the papers that deal with the Security and Privacy issues of Management in MCC are illustrated [26] [27] [28] [29] [30] [31] [32]. As we can realize there are several works in this field. More particular, in [26] the authors propose an entity-centric approach for an IDM model in Cloud environment. The proposed approach based on two aspects: a) active bundles, and b) anonymous identification. The active bundles include a payload of Personally Identifiable Information, privacy policies and a virtual machine that enforces the policies and additionally the active bundles use a set of protection mechanisms in order to protect themselves. As regard the anonymous identification, they use it with the aim to mediate interactions between the entity and the Cloud services using entity's privacy policies. Moreover, the authors present the main characteristics of the approach which are: a) independent of third party, b) provides minimum information to the Service Provider, and c) provides ability to use identity data on untrusted hosts. Then, the [27] demonstrates the implementation of a mobile system that enables electronic healthcare data storage, update and retrieval using Cloud Computing. The proposed mobile application based in Google's Android OS and offers management of patient health records and medical images. This system was evaluated with the use of Amazon's S3 cloud service. Finally, the authors summarize the details of the implementation and then present initial results of the system in practice. Moreover, the authors of [28] survey the MCC technology, which could help the general readers to have an overview of the MCC including the definition, the architecture, and the applications. Also, the [28] presents the issues, the existing solutions, and the recent approaches of the MCC technology. At the end, the authors discuss a number of future research directions of the MCC. Through the [29] the authors propose a multi-faceted Trust Management system architecture for a cloud computing marketplace, with the aim to support the customers in reliably identifying trustworthy cloud providers. The proposed system offers means to identify the trustworthy cloud providers in term of different attributes that assessed by multiple sources and roots of trust information.

Furthermore, the [30] presents a sort survey of MCC evolution and additionally explains how Cloud Computing and Mobile Devices could be combined with good terms for future opportunities, implications and legal issues for developing countries. In another research, the authors of [31] try to review the existing Distributed Application Processing Frameworks, also known as DAPFs, for SMDs in MCC domain. The main objective of [31] is to highlight issues and challenges to existing DAPFs in developing, implementing, and executing computational intensive mobile applications within MCC domain. Thus, through this work the authors propose a thematic taxonomy of the current DAPFs, and then they review current offloading frameworks by using thematic taxonomy, and analyze the implications and critical aspects of current offloading frameworks. Finally, the [31] puts forward open research issues in distributed application processing for MCC that remains to be addressed. Also, the [32] pro-poses a trust management approach by making an analysis of user behavioral patterns for a reliable Mobile Cloud Computing. So, the authors suggest a method in order to quantify a one-dimensional trusting relation count on the analysis of telephone call data from Mobile Cloud Environment. Subsequently, it is enhanced trustworthiness of data production, management, and overall application.

Finally, in [33] there is a proposal of an efficient algorithm for advanced scalable Media-based Smart Big Data, such as 3D and Ultra HEVC, on Intelligent CC systems. The proposed encoding algorithm of [33] exceeds the conventional HEVC standard which has been demonstrated by the performance evaluations.

Also, related works of other research groups have been studied. The [34] presents a survey on the BD and CC, with the importance to promote the research and development activities in the sector of the BD and the cloud computing. At the end, the [34] introduces a method for storing the data on cloud using the CloudSim simulation software.

Then, [35] shows an analysis that focuses on the two key concepts, BD and CC, and some of the issues and possibilities which are innate with the deployment of CC and BD services. Through this study is shown which security challenges is among the most prominent problem in CC and BD services. Finally, after there is a consideration about some of the problems related to BD and CC, a number of solutions that have been suggested in [35] towards improving the two key concepts that will go a long way in increasing the adoption rate of CC by organizations.

In [36] the authors surveys on the effects of data processing and analyzing big healthcare data on a CC environment. The [36] proposes the use of the Hadoop, which is a system that could process large amounts of data sets on distributed environments, and also it can be deployed on a CC environment to process the big healthcare data.

The authors in [37] propose an IoT-based security sys-tem on smart building scenarios. By this, they are integrating coherent data as fundamental components. The aim of the integration is to drive the building management and security behavior of indoor services accordingly. A holistic platform named City Explorer, which offers security and discovery, is the component in which the proposed system is manifested.

In [38] is illustrated an energy saving solution in buildings aiming to generate predictive models of energy consumption in buildings. Moreover, the authors in [38] use a building as a reference, for which they have one year’s unified data, in order to verify the proposed solution. At the end, the authors proposed strategies and control actions for energy saving in the building.

With the aim to take measurements about the temperature, the humidity, and the light in a building, the authors in [39] present an IoT-based sensing and monitoring system which is wirelessly connected. Also, in [39] there is a development of an Android application through which data is transmitted from the LabVIEW, to a “*smart*” mobile device through which data are monitored remotely.

In [40] the authors analyze the problem of imperfection in smart city data. Additionally, the authors point on the management of these types of data and also create an evidential database with the use of the evidence theory, with the aim to improve the efficiency of the smart city. Moreover, in this paper has been presented a special case of modeling imperfect data in the healthcare sector. Finally, a database which embraces both imperfect and perfect data was built up and the different imperfect aspects, in this database had been represented by the theory of beliefs and illustrated in this paper.

As an attractive service, has been characterized the data sharing service in [41]. As this paper informs us, the attribute based encryption (ABE) is widely discussed, and is the scheme on which the proposed scheme in this paper is based on. This scheme provides solutions for the resource constrained IoT-mobile devices in the clouds. The feasibility and efficiency of the scheme has been proved through performance analysis and experiments which confirm that the scheme is also protected of adaptively chosen ciphertext attacks.

The widely and continuous deployment and use of novel technologies usually leads to threats that come from internal and external factors. A research [42] which deals with the personal mobile data privacy of mobile users provides a protection scheme that is based on the "*Attribute-Based Access Control*" (ABAC) and the data self-deterministic schemes. The "*Attribute-based Semantic Access Control*" (A-SAC) algorithm and the "*Proactive De-terminative Access*" (PDA) algorithm have been used by the authors in [42] to support the proposed scheme. The benefits of the scheme are the constraining data accesses, the proactive prevention of the users' data threats on the cloud, and the increased level of secure sustainability.

Another region based approach that deals with the da-ta safety and the security mechanisms, in the healthcare sector this time, has been presented in [43]. The authors of this paper, through the blend of the RSA (Rivest-Shamir-Adleman) and the AES (Advanced Encryption Standard) algorithms, have been deployed a novel hybrid encryption scheme. The proposed scheme can protect the patients' personal information by concealment of them into a cover image. This image is characterized by high indistinctness, high capacity, and minimized distortion. The feasibility of the scheme is proved through the comparative analysis that was made between other state-of-the-art methods and the proposed one.

Moreover, the authors of [44] review the current re-search challenges and opportunities related to the development of secure and safe Intelligent Transport Systems (ITS) applications. Initially, they explore the architecture and main features of the ITS systems and also they survey the key enabling standards and projects. Likewise, the authors provide an analysis of a detailed ITS safety application case study and then evaluate in light of the European ETSI TC ITS standard.

Eventually, the [45] states that the Internet of Things could enable innovations that enhance the quality of life, nevertheless IoT generates unprecedented amounts of data that are difficult for traditional systems, Cloud Computing, and even the Edge Computing to handle. Consequently, Fog Computing is designed to overcome these limitations.

Additionally, there some "*key*" related research works that deal with the Security of the Machine Learning systems [46] [47] [48] [49] [50] [51]. Specifically, the [46] offers a framework for answering the major question "*Can machine learning be secure?*". The novel contributions of this work introduces: a) a taxonomy of different types of attacks on machine learning techniques and systems, b) a variety of defenses against those attacks, c) a discussion of ideas that are important to security for machine learning, d) an analytical model giving a lower bound on attacker's work function, and e) a list of open problems. The [47] focuses to offer a brief overview on the current work towards the emerging research problem of secure machine learning. Furthermore, the [47] presents a brief overview on secure machine learning and current progress on developing secure machine learning algorithms. Subsequently, the [48] presents taxonomy which identifying and analyzing attacks against machine learning systems. In addition to this, the authors of [48] show how these classes influence the costs for the attacker and defender, and we give a formal structure defining their interaction. At the end, this work presents a discussion of how the proposed taxonomy suggests new lines of defenses. The authors of [49] design a novel, communication-efficient, failure-robust protocol for secure aggregation of high-dimensional data. Their proposed protocol allows a server to compute the sum of large, user-held data vectors from mobile devices in a secure manner, and can be used, for example, in a federated learning setting, to aggregate user-provided model updates for a deep neural network. Through their work, the authors of [49] prove the security of their protocol in the honest-but-curious and active adversary settings, and show that security is maintained even if an arbitrarily chosen subset of users drop out at any time. Also, the authors evaluate the efficiency of their protocol and show, by complexity analysis and a concrete implementation, that its runtime and communication overhead remain low even on large data sets and client pools. In [50] the authors rely upon a previously-proposed attack framework to categorize potential attack scenarios against learning-based malware detection tools, by modeling attackers with different skills and capabilities. Then, the authors of [50] try defining and implementing a set of corresponding evasion attacks to thoroughly assess the security of Drebin, an Android malware detector. As a result, the main contribution of this work is the proposal of a simple and scalable secure-learning paradigm that mitigates the impact of evasion attacks, while only slightly worsening the detection rate in the absence of attack. At the end, the authors argue that their secure-learning approach can also be readily applied to other malware detection tasks. Finally, the authors of [51] propose a DSQML protocol in which the client can classify two-dimensional vectors to different clusters, resorting to a remote small-scale photon quantum computation processor. The proposed protocol is secure without leaking any relevant information. Regarding the principle, the proposed protocol can be used to classify high dimensional vectors and may provide a new viewpoint and application for future “*Big Data*”.

7.3 Big Data

Big Data is the concept of data where it is difficult to gather, store, handle and process with classic tools and technologies. Over the last two decades, Big Data in the industry has grown enormously in various sectors and is growing exponentially. In 2011, the volume of data generated in the world was 1.8ZB and this will double every two years in the near future [4] [5].

The concept of large data has been defined by the 3V model from Lenay [52] as: “*high volume, high speed and a wide variety of information items that require efficient and innovative forms of information processing for improved insight and decision making*” [4] [11].

In 2012, Gartner [52] updated the definition as follows: "*Big data is high-intensity, high-speed, and/or high-variety of information items that require new forms of processing to enable enhanced decision making, discovery of insight optimization of processing*". The TechAmerica Foundation [53] defines the large data as follows: "*Big data is a term describing high-speed, complex and variable high-volume data that requires advanced technologies and techniques to enable capture, storage, distribution, management and analysis of information*".

7.3.1 Predictive model of Big Data’s 5V

For predicting Big Data’s 5V, a real-time system is pro-posed that initially filters data from unreliable sources (honesty) and distinguishes the variety of data using the Bloom filter [54]. It then uses the Kalman filter to estimate the volume and speed of each data variety that arrives in the system, the data variability is incorporated while the volume and speed are estimated. Kalman filter could be characterized as better filter than the other filters as it can be easily adapted to provide impartial estimates across a wide range of data streams even when the fluctuation is high. It is an effective retrospective filter, a mathematical toolkit capable of dynamically predicting future trends from incoming currents from sensor measurements with noise [21]. The Bloom filter is a probabilistic data structure that is used to filter data that does not belong to a set. Data streams consider it to be mainly: text, audio, video and video data [54].

7.3.2 Big Data Analytics

The creation of heterogeneous data from different physical devices requires quick real-time analysis. Incomplete data is a problem for real-time analysis, so we need algorithms that pre-process the data before analysis.

As production data continues and grows, the way in which Big Data can expand and follow this evolution is a challenge [3] [4] [21] [33].

One of the most important benefits of the Internet of Things Technology is the creation of an unprecedented amount of data. Storing, holding and completing data becomes critical. The internet consumes up to 5% of the total energy produced today and with these requirements, it will certainly increase even more. As a result, centralized and centralized data centers ensure both energy efficiency and reliability. The data must be stored and used intelligently for intelligent monitoring and activation. It is important to develop artificial intelligence algorithms that can collect or distribute depending on the current needs. New fusion algorithms need to be developed to understand the data collected. The modern non-linear, time machine learning methods based on evolutionary algorithms, genetic algorithms, neural networks and other artificial intelligence techniques needed for automated decision making. These systems present features such as interoperability, integration and adaptive communications. They also have a modular architecture both in terms of hardware design and software development and are usually suitable for IoT applications. What is needed is the existence of a central infrastructure to support storage and analysis. This makes the IoT intermediate software level and there are many challenges that are discussed below. Since 2012, the storage solutions based on Cloud are becoming increasingly popular in the coming years under analysis platforms based on the Cloud and data visualization platforms collected [3] [5] [12] [21].

Data analysis is the process of using algorithms that are executed on powerful platforms to discover hidden capabilities in large data such as hidden patterns or unknown associations, for example, the extraction of useful knowledge and their image [55]. This is done in the wording of the case, often based on conclusions gathered from the experience and the discovery of correlations between the variables [56]. According to Rajaraman et al [56], there are four types of data analysis:

Descriptive Analysis: This deals with what has happened in the past and presents in a readily understandable form the data such as diagrams, graphs, pie charts, maps, spreadsheets, etc., the display gives an insight into what the data imply. A typical example is the presentation of population census data that classifies the population in a country by gender, age, education, income, etc [56].

Predictive Analysis: It draws conclusions from the available data to say what is expected to happen in the near future. The tools used to collect data are time series analysis using statistical methods, neural networks, and engineering learning algorithms. An important use of predictive analysis is in marketing that understands the needs and preferences of customers [56].

Exploratory Analysis: Finds unexpected relationships between parameters in large data collections. Collecting data from various sources and analyzing them provides additional opportunities for new ideas and random discoveries. One of the most important applications is to discover patterns in customer behavior from the feedback they get from tweets, blogs, Facebook, emails to allow companies to predict customer actions such as renewing subscription to the magazine, changing a mobile phone service provider, canceling a hotel reservation, and so on [56].

Regulatory Analysis: It identifies, based on the data gathered, opportunities to optimize solutions to existing problems, ie tells us what needs to be done to achieve a goal. One of the common uses is the pricing of airlines based on data from travel models such as: popular destinations and destinations, major events, holidays etc. to maximize profit [56].

Moreover, Alexandrov et al. [57] present Stratosphere, which is an open source software for parallel data analysis. In addition, Kwon et al. [58] propose a research model to explain the intent to buy large analytical data, mainly from the theoretical approaches to data quality management and user experience.

7.3.3 Big Data Security Issues

New challenges and standards developed and created in data security issues through the development and the use of BD technology. This creates a growing need for further research on security technologies in order to be able to handhold the large amount of data and to ensure effective. Technologies for securing data are slow when applied to huge amounts of data [3] [12] [21] [33].

Algorithm	Key length	Megabytes processed	Block size	Rounds	Time Taken	MB per Second
3-DES	56, 112 or 168 bits	128	64 bits	48	6,159	20,783
AES	128, 192 or 256 bits	256	128 bits	10, 12, or 14	4,196	61,010
RSA	1024-4096 bits	300	512 bits	1	1175,7826	10,900

Table 7.1: Encryption rates of popular algorithms

Regarding the Table 7.1 we can conclude that even the most efficient algorithms give an encryption rate of 64.3MB/s. So, in the sector of BD technology, in which the need of large amounts of data need to be transferred we can see a significant bottle neck for encryption such large amounts data. This is detrimental to the nature of BD which has real time processing and results.

7.3.4 Big Data on Cloud System Scenario

Among all types of data in the cloud storage, large-scale data has occupied a significant part due to the explosive sharing on social networks and additionally video-on-demand services for movies, TV programs, etc. Moreover, to support users with various bandwidth requirements and device resolutions and full interactive playback in large-scale data demand, usually various versions at different bitrates are generated [3] [12] [21] [33] [59] [60] [61].

Schemes for large-scale data, named as Big Data, have shown good performances in cloud storage under different configurations. However, these codes treat all files as general data, in which one unrecoverable error will lead to permanent loss of the whole file. They do not consider the features of specific data types.

The Cloud Computing should provide its services with specific functions so that the IoT linked to it, can support the smart city's turn. The Big Data, or large scale data, as it described in the international literature is defined as the large quantity data that specific scenarios described, relate to the whole activity of the city.

In this work, we propose Cloud-based system for BD used and transmitted through an IoT network.

7.4. Internet of Things

The IoT could be characterized as "*a network of devices that transmits, shares, and uses data from the physical environment to provide services to individuals, corporations, and society*" [1] [8] [12], which already defined in the Introduction Section. Also, IoT has multiple applications in health, transport, environment, energy or types of devices such as sensors, devices worn/carried (wearable), watch, glasses, home automation (domotics).

7.4.1 Advantages of the data

Chances where the streaming data will produce novel markets with the aim to inspire positive change or to intensify existing services are examined by businesses. Some examples of fields that are at the heart of these developments are listed below [62]:

a) ***IoT(a)***: Smart solution in the bucket of transport: With this could achieve better solutions in transportation sector with the aim to provide a better way of living.

b) ***IoT(b)***: Smart power grids incorporating more renewable: With this the system reliability could be achieved and also it could be reduced the charges consumers, thus providing cheaper electricity.

c) ***IoT(c)***: Remote monitoring of patients: With this we could achieve a system which offers remote monitoring of patients. This system could offer a better and well-managed healthcare system by improving the quality of services, increasing the number of people served, and saving money.

d) ***IoT(d)***: Sensors in homes and airports: With this we could achieve safer places such as airports and houses, by establishing a number of sensors in the field.

e) ***IoT(e)***: Engine monitoring sensors that detect & predict maintenance issues: With this we detect and predict maintenance issues, improve inventory replenishment, and even define priorities in scheduling maintenance work, repairs, and regional operations.

7.4.2 IoT Data

IoT is an example of networking where cyber-physical systems consisting of automatic sensors, actuators and embedded systems are associated with the physical world including the human being for real-time support, security, personality and high-level performance [1] [8]. IoT has great potential in manufacturing [3] [12].

Cyber-physical systems, smart devices, industrial instruments, sensors, actuators, OPC Server are examples of IoT devices that produce heterogeneous data.

Data collected from the following IoT technologies play an important role:

1. **Radio Frequency Identification (RFID):** RFID technology uses electromagnetic fields for data transfer as well as automatic object detection [22]. It consists of tags and readers. Each device has a unique RFID tag. The reader detects objects by reading labels. Storing and managing RFID data is a challenge for large businesses as only certain items and products have RFID tags.

2. **Wireless Sensor Network (WSN):** WSN is a network of distributed autonomous nodes connected to other nodes via wireless sensors in a limited environment [2] [22]. The sensor node is self-organizing and connected to other nodes to transmit its data back to the central grid. Some nodes have the ability to control actuators (physical devices) in the sense of automation. WSNs contain all the node information that have sensors and actuators to communicate and transfer their commands [3] [6].

3. **Cloud Computing:** Today, storage, computing power, infrastructure, platforms and software can only be offered as a service by paying only as we use them. Infrastructure as a Service (IaaS), Platform as Service (PaaS) and Software as a Service (SaaS) are the three main cloud computing models. The architecture of IoT Cloud computing plays an important role for IoT data. They can be stored in Cloud and accessible from anywhere and anyone using an Internet Browser or software [11].

4. **Industrial Internet:** The Industrial Internet, also known as the Industrial Internet of Things (IIoT), is the Internet of Things (IoT) only for industries. Smart Machines Link Industrial World both internally and externally facilitating communication using advanced hardware and software [4].

7.4.3 Security

The security of IoT systems is a field of strives concerned with safeguarding connected devices and networks in the IoT. The IoT involves the growing pervasiveness of objects and the entities provided with unique identifiers and the ability to automatically transmit data through a network. The major impact of the increased use of IoT communication came from computing devices and embedded sensor systems which used in industrial machine-to-machine (M2M) communication, and technologies such as smart energy grids, home and building automation, vehicle to vehicle communication and wearable computing devices [2] [22] [63] [64].

The huge issue is that security has not always been considered in product design due to the idea of networking appliances and other objects were relatively new. Aiming to improve security and privacy issues, an IoT device that needs to be directly accessible through the Internet should be portioned into its own network and has limited network access. The network portion should be monitored in order to identify the potential abnormal traffic, and if there is any problem, action should be taken [2] [22] [63] [64] [65].

In the sector of IoT technology there are System models. A wireless network model with a source-destination pair, N trusted relays and J eavesdroppers $(J \leq 1)$ are considered. Suppose that the global CSE is available. The eavesdropper channel, source encoding schemes, decoding models and accommodative protocol are admitted to be public, only source message is assumed to be confidential. In this work, the discussion is limited to two main accommodative models: Decode-and-Forward (DF) and Amplify-and-Forward (AF) [65] [66] [67].

Decode-and-forward (DF)

Two are the main stages in DF model. In Stage 1, the source broadcasts its encoded symbols to its trusted relays using the first transmission slot. When the symbol x transmitted, the received signals at the N relays are given by (1),

$$y_r = \sqrt{P_s} h_{SR}^* x + n_r \quad (1)$$

where P_s is the transmit power of source and n_r is the noise vector at relays [66].

In Stage 2, all the trusted relays that successfully de-code the message, re-encode the message and accommodative transmit the re-encoded symbols to the destination by using the second transmission slot. Each relay transmits a weighted version of the re-encoded symbol. When transmitting the symbol $\tilde{x}$, the received signal at the destination is given by (2),

$$y_d = h_{RD}^{\dagger} w \tilde{x} + n_d \quad (2)$$

while the received signal at the listeners is expressed in vector form as (3),

$$y_e = H_{RE}^{\dagger} w \tilde{x} + n_e \quad (3)$$

The transmit power budget for Stage 2 is considered to be P - P_s where P is the total power for transmitting one symbol and Ps is the transmit power of source [66].

<u>Amplify-and-forward (AF)</u>

At the other hand, the AF model is additionally a two-stage model such as the DF model. The Stage 1 is similar for both AF and DF models, except that the transmit power can be different. The trusted relays forward the signals that are received during Stage 1 to the destination, using the second transmission slot in Stage 2. That is, each relay transmits a weighted version of the noisy signal that they received during Stage 1. The transmitted signals of all relays are denoted by the product of $diag\{w\}y_r$, where w is the weight vector and y_r is given by (1). The received signal at the destination is given by [66],

$$y_d = \sqrt{P_s}\, h_{RD}^{\dagger} diag\{w\} h_{SR}^{*} x + h_{RD}^{\dagger} diag\{w\} n_r + n_d \quad (4)$$

The received signals at the listeners, in a vector form, is denoted by [49],

$$y_e = \sqrt{P_s}\, H_{RE}^{\dagger} diag\{w\} h_{SR}^{*} x + H_{RE}^{\dagger} diag\{w\} n_r + n_e \quad (5)$$

Also, another security challenge in IoT is the encryptions algorithm. The RSA algorithm, which is the most commonly used public key algorithm in the Internet, and it can be used in sensor networks by establishing a Trusted Platform Module (TPM), which costs less than 5% of a common sensor node [67]. So, the memory has been measured for a fully authenticated handshake with 2048-bit RSA keys. This type of handshake has the largest memory requirements since it needs more code and buffer space for the client’s Certificate and Certificate-Verify messages. The memory increased its use because the code basically contains hundreds of statements form buffer[x] = 0xff. The use of this encryption algorithm in IoT’s security could offer better communication privacy in its functionality.

7.5 Cloud Computing

CC offers abilities and functions such as computing, storage, services, and applications over the Internet. In general, to render smartphones energy efficient and computationally capable, major changes to the hardware and software levels are required. This causes the cooperation of developers and manufacturers [68].

7.5.1 Features

As all technologies, so the CC technology has a number of characteristics which determine its operation. These characteristics are represented and outlined below.

***CC(a):** Storage over Internet*

Storage over Internet can be defined as "*a technology framework that uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks to link servers and storage devices and to facilitate storage solution deployment*" [69] [70].

***CC(b):** Service over Internet*

The Service over Internet has as major objective is to "*help customers all over the world in order to transform aspirations into achievements by harnessing the Internet's efficiency, speed and ubiquity*" [69] [70].

***CC(c):** Applications over Internet*

Cloud Applications, or as scientific known as Applications over Internet, are the programs which have been written to do the job of a current manual task, or virtually anything, and which perform their job on the server through an internet connection [69] [70].

***CC(d):** Energy Efficiency*

Energy Efficiency could be defined as "*a way of managing and restraining the growth in energy consumption*" [69] [70]. By delivering more services for the same energy input or for the same services for less energy input may be something more energy efficient [69] [70].

***CC(e):** Computationally Capable*

The services of computational clouds are leveraging the computationally concentrated and ubiquitous mobile applications which have been enabled by the technology of MCC. Thus, a system can be considered as computationally capable when it meets the requirements to offer us the results we want, by making the right calculations [69] [70].

7.5.2 Security on Cloud Computing

CC security is an evolving sub-domain of computer security, network security and information security. It makes an allusion to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of CC.

CC technology offers through its storage solutions to users and industries various capabilities with the aim to store and process their data in third-party data centers [71]. Thus, by aiming to offer secure communication through the network, encryption algorithm plays a vital role. As regards the researches that have been made, an important encryption technique is the Symmetric Key Encryption. In

Symmetric key encryption, only one key is used to encrypt and decrypt the data. In this encryption technique the most used algorithm is the AES [72] [73].

AES (Advanced Encryption Standard) is the newest encryption standard and the more reliable, recommended by NIST to replace DES algorithm. The only effective scenario of attacking in AES is the Brute force attack, in which the attacker tries to test all the characters combinations to unlock the encryption. AES encryption model is fast and flexible, and in addition, it can be implemented on different platforms [74]. Bellow, a sample-part of the AES encryption algorithm is represented.

Algorithm: sample of AES

```
Cipher(byte[] input, byte[] output)
{
    byte[4,4] State;
    copy input[] into State[] AddRoundKey
    for (round = 1; round < Nr-1; ++round)
    {
        SubBytes ShiftRows MixColumns AddRoundKey
    }
    SubBytes ShiftRows AddRoundKey
    copy State[] to output[]
}
```

AES algorithm characterized as better and safer than other algorithms for a number of reasons, which is follows [75]:

- It performs consistently well in both hardware and software platforms under a wide range of environments. These include 8-bit and 64-bit platforms and DSP's.
- Its inherent parallelism facilitates efficient use of processor resources resulting in very good soft-ware performance.
- This algorithm has speedy key setup time and good key agility.
- It requires less memory for implementation, making it suitable for limited-space environments.
- The structure has good potential for benefiting from instruction-level parallelism.
- There are no serious weak keys in AES.
- It supports any block sizes and key sizes that are multiples of 32 (greater than 128-bits).
- Statistical analysis of the cipher text has not been possible even after using huge number of test cases.
- No differential and linear cryptanalysis attacks have been yet proved on AES.

7.5.3 Cloud Computing trade offs

Cloud Computing has some disadvantages-limitations which should be eliminated over the years in order to achieve a better and more ideal use. Some businesses and especially the smaller ones need to be aware of these limitations before going in for this technology.

CC(l-a): *Security*

One major issue of the Mobile Cloud Computing is the security issue. Before someone adopts this technology, they should know that all the company's sensitive information would be surrender to a third-party Cloud service provider. This could potentially put the company in great risk. Hence, someone must be absolutely sure that they would choose the most reliable service provider, who will keep the information completely safe [11] [76] [77].

CC(l-b): *Connectivity*

Internet connection is critical to Cloud Computing. Thus, the user should be certain that there is a good result before opting for these services. Since someone owes a mobile device which is connected to the internet has become the norm in the wireless world of today, Cloud Computing has a very large potential user base [11] [78].

CC(l-c): *Performance*

Another major concern of the Cloud Computing pertains to its performance. Some users feel performance is not as good as in native applications. Thus, checking with one service provider and understanding their track record is advisable [11] [79] [80].

CC(l-d): *Latency (Delay)*

In Cloud Computing, latency (sometimes referred as turnaround time) is defined as the time involved in offloading the computation and getting back the results from the nearby infrastructure or cloud [11] [15].

CC(l-e): *Privacy*

Data privacy is important and is one of the main bottlenecks that restrict consumers from adopting Cloud Computing. Therefore, to gain consumers trust in the Cloud, the application models must support application development with privacy protection, and implicit authentication mechanisms [11] [77] [81].

7.6 IoT & Cloud Computing Integration

Moreover, a new generation of services, count on the concept of the "*cloud computing*", has made its appearance in the last few years with the purpose of offering access to services and the data from any place and at any time [82]. CC is a technology that can be set as a base technology in the use of IoT [83].

A number of the major characteristics of the CC technology which relate to the features of IoT are: a) Storage over Internet, b) Service over Internet, c) Applications over internet, d) Energy efficiency and e) Computationally capable. Tables 7.2

presents the features of CC regarding the accessibility of this technology provides when combined with the characteristics of IoT [82] [83].

Internet of Things characteristics	*CC(a)*	*CC(b)*	*CC(c)*	*CC(d)*	*CC(e)*
IoT(a)	**X**	**X**	**X**		**X**
IoT(b)	**X**	**X**		**X**	**X**
IoT(c)		**X**	**X**		**X**
IoT(d)	**X**	**X**	**X**	**X**	**X**
IoT(e)		**X**	**X**	**X**	**X**

Table 7.2: Contributions of Cloud Computing in Internet of Things

Table 7.2 represents the characteristics of CC technology regarding the suitableness of this technology provides. Furthermore, it enumerates the major features of the IoT technology. The main objective of Table 7.2 is to show which of the specific characteristics of CC technology, related more and improve the functionality of the characteristics of IoT technology. As we can observe from Table 7.2, the characteristic of IoT which affected more by the characteristics of CC is "*Sensors in homes and airports*". Regarding the CC, the feature which affected more are "*Service over Internet*" and "*Computationally capable*". As a general conclusion, we can observe that those two technologies contribute more each other in many of their features.

7.6.1 Security issues in IoT and Cloud Computing integration

There is a rapid and self-sufficient evolution taking into account the two technologies of IoT and CC. Initially, the virtually unlimited capabilities and resources of CC with aim to remunerate its technological constrains, such as processing, storage and communication, could be a beneficial scenario for the IoT technology. In many cases, CC can offer the transitional layer between the things and the applications, hiding all the complexity and functionalities which are necessary to implement the latter [84].

Through the integration of IoT and CC could be observed that CC can fill some gaps of IoT such the limited storage and applications over internet. In the other hand, IoT can also fill some gaps of CC such the major problem of limited scope. Count on motivations such those referred previously and the important issue of security in both technologies we can consider some drivers for the integration. The security issue of this integration has a serious problem. When critical IoT applications move towards the CC technology, concerns arise due to the lack of trust in the service provider or the knowledge about service level agreements (SLAs) and knowledge about the physical location of data. Consequently, new challenges require particular attention as mentioned in surveys [85] [86]. Moreover, public key cryptography could not be applied at all layers due to the computing power constraints imposed by the things

[85]. These are examples of topics that are currently under examination in order to tackle the big challenge of security and privacy in CC and IoT integration [84].

Subsequently, some challenges about the security problem in the integration of those technologies are listed below [84].

a) *Heterogeneity*: A big challenge in CC and IoT integration is related to the wide heterogeneity of devices, operating systems, platforms, and services available and possibly used for new or improved applications [87].

b) *Performance*: Often CC and IoT integration's applications introduce particular performance and QoS requirements at several levels and in some specific scenarios meeting requirements might not be easily achievable [88].

c) *Reliability*: When CC and IoT integration is adopted for mission-critical applications, reliability concerns typically arise [89].

d) *Big Data*: With an estimated number of 50 billion devices that will be networked by 2020, particular attention must be paid to transportation, storage, access, and processing of the large amount of data they will produce [90].

e) *Monitoring*: This is an essential activity in CC environments for capacity planning, for managing resources, SLAs, performance and security, and for troubleshooting [91].

IoT & Cloud Computing security challenges	*Internet of Things*	*Cloud Computing*
Heterogeneity		**X**
Performance	**X**	**X**
Reliability	**X**	
Big Data	**X**	**X**
Monitoring	**X**	

Table 7.3: Affects of IoT & Cloud Computing security challenges

Table 7.3 shows the two technologies that we survey in this work and the challenges of their integration that arising from our study. These challenges are related to the security problem in the integration of two aforementioned technologies and they listed in detailed in subsection 6.1 (A Security issues in IoT and Cloud Computing integration). As we can observe from Table 7.3, the both technologies have two common main challenges of their integration which are Performance and Big Data. Additionally, we can observe that IoT technology is related to more challenges (4) than the CC technology (3).

7.6.2 Big Data based on Cloud Server

In order to combine BD technology with CC technology and to achieve a beneficial operation of BD in Cloud environment we have to study the relation of their basic features [3] [12] [22] [64].

Initially, we have to define which are the basic features of BD, which are widely known as the 5 Vs of Big Data. In particular the 5 Vs of BD are: 1) *Volume*: the vast amounts of data created every second, 2) *Velocity*: the speed at which new data is created and the speed at which data moves around, 3) *Variety*: the different types of data we can now use. In the past we focused on structured data that neatly fits into tables or relational databases, such as financial data, 4) *Veracity*: the messiness or trustworthiness of the data, 5) *Value*: all well and good having access to big data but unless we can turn it into value it is useless [22] [64].

Big Data Features *Cloud Computing Features*	**Volume**	**Velocity**	**Variety**	**Veracity**	**Value**
Storage over Internet		X		X	X
Service over Internet	X		X	X	X
Applications over Internet	X	X	X	X	X
Energy Efficiency	X	X			
Computational Capable		X	X		X

Table 7.4: Correlation of BD and CC characteristics

Table 7.4 demonstrates the basic features of BD (5 Vs) and how they are contributed by the major features of CC. As we can observe, there are two the key features of BD technology which contributes more with the characteristics of CC technology are *Velocity* and *Value*. *Velocity* and *Value* contribute four from the five key features of CC. Also, another thing that we can observe from Table 7.4 is that the feature *Applications over Internet* contributed from all the key features of BD.

7.6.3 Proposed Efficient IoT and Cloud Computing Security Model

As we can infer, by taking advantage of the reasons which AES algorithm offers better secure in CC and the two models that give benefits in security problems in IoT we can propose a novel method that uses those benefits with the aim to improve the security and privacy problems in the integration of two technologies.

The AES algorithm offers the ability to have speed key setup time a good key agility. So, if we use this algorithm in the functionality of DF model, we could have a trusted relay method with an encryption of a speed key setup. Therefore, instead the trust relay use that DF and AF methods offer we can seize additionally there no serious weak keys in AES and so we could have a beneficial security use of the encryption in the integrated new model. Moreover, we can take advantage the less memory which AES needs for implementation that makes it for restricted-space environments. So, we can seize the transmit power that the AF model offers and as a result we can have a better and more trusted transmission. In the way of transmission, when the symbol transmitted with the use of DF model, the received signal at destination is given by the equation (2), which mentioned in previous section.

With this proposed model we can extend the advances of IoT and CC, by developing a highly innovative and scalable service platform to enable secure and privacy services. Through this research we can propose the following part of algorithm which extends the security advances of both technologies. As a proposal of this work could be this part of pseudocode algorithm which uses the original key consists of 128 bits/16 bytes which are represented as a 6x6 matrix, represented bellow.

Algorithm 1: pseudocode

```
input -> byte[]
byte[] + R.Key -> state[]
for 6 to 66
  W[i-1] -> T
  if i mod 6 = 0
              rotate T + 6
      W[i-6] / T -> W[i]
      R.Key+1
      i+1 -> i
Row +1 -> Row
state[] -> output[]
```

Algorithm 1 represents the procedure implementing in the server aiming to achieve better results of securing the data transmitted. Moreover, this procedure could be achieved in a limited number of loops of the algorithm. The algorithm takes as input data the transmitted signal and then with the use of AES algorithm and the key generated tries to decrypt the data by using the original key consists of 128 bits/16 bytes which are represented as a 6x6 matrix. Through this procedure we could achieve the less of loops of the algorithm and in addition to this we can achieve a more secure data decryption/encryption system for transmitting the data through the network.

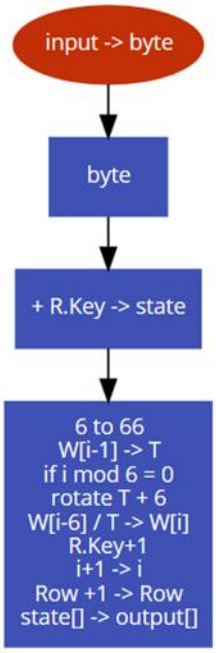


Figure 7.1: Flowchart of the proposed the procedure implementation

Figure 7.1 shows the proposed pseudocode representation through a flowchart.

7.6.4 Experimental Results

Considering the benefits of the security models and algorithms of IoT and CC technologies we can observe that we can have a beneficial use of integration those two technologies. Instead of the wide use of IoT we can take advantage that CC security through the AES algorithm performs consistently well in both hardware and software platforms under a wide range of environments. This use could be possible for all type of platforms and DSPs. Furthermore, the novel integrated technology could has good potential for benefiting from instruction-level parallelism and will support any type of block sizes and key sizes that are multiples of 32 and used both of IoT and CC. Also, each transmitted signal through the new technology can transmitted as a relay and trusted signal with a weighted version of the re-encoded symbol.

Through this integration we can achieve some useful functions, i.e. we can use the Cloud-based IoT service with the aim to connect sensors and additionally made them capable to share the sensor readings with others, reducing the security issues. Furthermore, another useful operation is that we can use the HTTP protocol with the aim to send data between IoT things and the CC applications. Moreover, some of the key advantages and challenges that can be defined from this integration are: 1) Both the physical hardware manufacturing resource and software manufacturing can be intelligently perceived and connected into the wider networks with the support of IoT technologies. 2) The collected information and data can be communicated and transmitted between M2M under the support of specific IoT technologies. 3) The collected and transmitted information can be processed and computed according to particular requirements under the support of different CC service, and some useful data and decision information can be intelligently generated and obtained.

AES Characteristics	*Internet of Things*	*Cloud Computing*	*IoT & CC integration*
Key length	X	X	X
Rounds		X	X
Certifications	X	X	X
Speed	X		X

Table 7.5: AES contribution in IoT and Cloud Computing

The Tables 7.5 exhibiting the key features of the two encryption algorithm that used with the aim to achieve integration of the technologies of IoT and CC concerning the security problem. Table 7.5 presents which of the key features of AES encryption algorithm contributes both IoT and CC technologies, and at the end how completely contributes the integration model of IoT and CC.

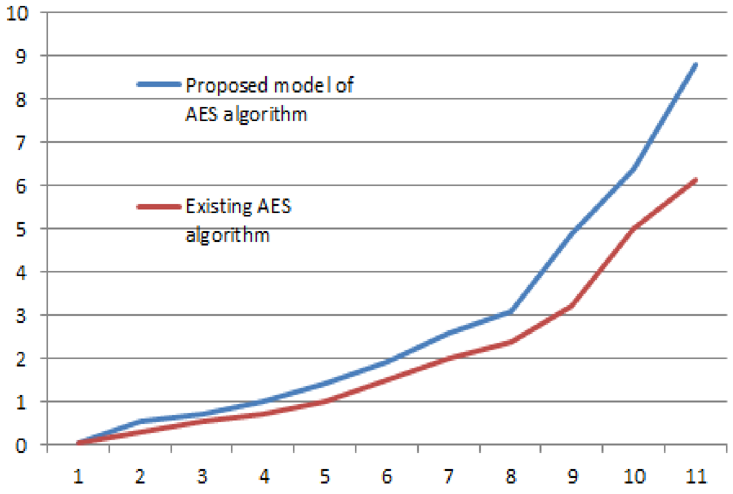


Figure 7.2: Security level of encryption algorithms of measurement used for the study of AES model algorithm

Figure 7.2 shows, the measurements that have been through time. As we can observe by this figure the more often is the combined use of the algorithms, the higher level of security of the data usage we get every time. The upper line represents our proposed model of AES algorithm and the other (down line) represents the existing AES algorithm.

Big Data Features	**Volume**	**Velocity**	**Variety**	**Veracity**	**Value**
IoT & CC integration model	**X**	**X**	**X**	**X**	**X**

Table 7.6: Correlation of BD characteristics with IoT & CC integration model

The Tables 7.6 exhibits the key features of BD and which of those characteristics could be contributed by the integration method of the technologies IoT and CC concerning the security problem. Table 7.6 presents that all the characteristics of BD contributed by the integration model of IoT and CC technologies.

7.7 Chapter Summary

The CC technology provides a number of possibilities, but additionally places several limitations as well. Cloud Computing refers to an infrastructure where both the data storage and the data processing happen outside of the mobile device. Also, the IoT is a new technology which is growing rapidly in the field of telecommunications, and especially in the modern sector of wireless telecommunications.

The main objective of the interaction and cooperation between things and objects sent through the wireless networks is to fulfil the objective set to them as a combined entity, with the aim to achieve a better environment for the use of Big Data. In addition, based on the technology of wireless networks, both the technologies of CC and IoT develop rapidly. In this work, we present a survey of IoT and CC with a focus on the security problems of both technologies. Particularly, we combine the two aforementioned technologies with the aim to examine the familiar characteristics, and with the aim to discover the benefits of their integration in order to secure the use and the transmission of Big Data.

At the end, the security challenges of the integration of IoT and CC were surveyed through the proposed algorithm model, and additionally there is a presentation of how the two encryption algorithms which were used con-tributes in the integration of IoT and CC as base technologies for Big Data. This and additionally the security challenges that surveyed in this work can be the domain of future research on the integration of those two technologies.

7.8 Chapter References

[2] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, December 2016.

[3] A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, B. B. Gupta, “Efficient IoT-based sensor BIG Data collection-processing and analysis in Smart Buildings”, Future Generation Computer Systems, vol. 82, pp. 349-357, May 2018.

[4] M. Hilbert, P. López, “The World’s Technological Capacity to Store, Communicate, and Compute Information”, Science, vol. 332, issue: 6025, pp. 60–65. 2011.

[5] Z. Fu et al, "Enabling Personalized Search over Encrypted Out-sourced Data with Efficiency Improvement", IEEE Transactions on Parallel and Distributed Systems, 2015.

[7] C. Stergiou, K. E. Psannis, “Efficient and Secure Big Data deliv-ery in Cloud Computing”, Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017.

[8] C. Stergiou, K. E. Psannis, A. P. Plageras, G. Kokkonis, Y. Ishibashi, “Architecture for Security in IoT Environments”, in Proceedings of 26th IEEE International Symposium on Industrial Electronics, 19-21 June 2017, Edinburgh, Scotland, UK.

[11] C. Stergiou, K. E. Psannis, "Recent advances delivered by Mo-bile Cloud Computing and Internet of Things for Big Data applications: a survey", Wiley, International Journal of Network Management, pp. 1-12, May 2016.

[12] C. Stergiou, K. E. Psannis, A. P. Plageras, Y. Ishibashi, B.-G. Kim, “Algorithms for efficient digital media transmission over IoT and cloud networking”, Journal of Multimedia Information System, vol. 5, no. 1, pp. 1-10, March 2018.

[15] J. Li, L. Huang, Y. Zhou, S. He, Z. Ming ,"Computation partitioning for mobile cloud computing in big data environment", IEEE Transactions on Indus-trial Informatics, Vol. 11 January 2017.

[21] C. Stergiou, K. E. Psannis, “Algorithms for Big Data in Advanced Communication Systems and Cloud Computing”, in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece.

[22] A. P. Plageras, C. Stergiou, K. E. Psannis, Byung-Gyu Kim, Brij Gupta, Y. Ishibashi, “Solutions for Inter-connectivity and Security in a Smart Hospital Building”, in Proceedings of 15th IEEE International Conference on Industrial Informatics (INDIN 2017), 24-26 July 2017, Emden, Germany.

[24] A. P. Plageras, C. Stergiou, K. E. Psannis, G. Kokkonis, Y. Ishi-bashi, Byung-Gyu Kim, Brij Gupta, ”Efficient Large-Scale Medical Data (eHealth Big Data) Analytics in Internet of Things”, in Proceedings of 19th IEEE International Conference on Business Informatics (CBI’17), International Workshop on the Internet of Things and Smart Services (ITSS2017), 24-26 July 2017, Thessaloniki, Greece.

[25] C. Stergiou, K. E. Psannis, A. P. Plageras, T. Xifilidis, B. B. Gupta, “Security and Privacy of Big Data for Social Networking Services in Cloud”, in Proceedings of IEEE conference on Computer Communications (IEEE INFOCOM 2018), 15-20 April 2018, Honolulu, HI, USA.

[26] P. Angin, B. Bhargava, R. Ranchal, N. Singh, M. Linderman, “An Entity-centric Approach for Privacy and Identity Management in Cloud Computing”, in Proceedings of 29th IEEE International Symposium on Reliable Distributed Systems, 31 October-3 November 2010, New Delhi, India. [DOI: 10.1109/SRDS.2010.28]

[27] C. Doukas, T. Pliakas, I. Maglogiannis, “Mobile Healthcare Information Management utilizing Cloud Computing and Android OS”, in Proceedings of 32nd Annual International Conference of the IEEE EMBS 2010, 31 August-4 September 2010, Buenos Aires, Argentina.

[28] H. T. Dinh, C. Lee, D. Niyato, P. Wang, “A survey of mobile cloud computing: architecture, applications, and approaches”, Wireless Communications and Mobile Computing, vol. 13, pp. 1587-1611, October 2011. [DOI: 10.1002/wcm.1203]

[29] S. M. Habib, S. Ries, M. Muhlhauser, “Towards a Trust Management Sys-tem for Cloud Computing”, in Proceedings of IEEE International Joint Conference TrustCom-11/IEEE ICESS-11/FCST-11, 16-18 November 2011, Changsha, China.

[30] M. R. Prasad, J. Gyani, P.R. K. Murti, “Mobile Cloud Computing: Implications and Challenges”, Journal of Information Engineering and Applications, vol. 2, no. 7, pp. 7-15, October 2012.

[31] M. Shiraz, A. Gani, R. H. Khokhar, R. Buyya, “A Review on Distributed Application Processing Frameworks in Smart Mobile Devices for Mobile Cloud Computing”, IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1294-1313, November 2012.

[32] M. Kim, S. O. Park, “Trust management on user behavioral patterns for a mobile cloud computing”, Springer, Cluster Computing, vol. 16, issue 4, pp. 725-731, December 2013. [DOI: 10.1007/s10586-013-0248-9]

[33] C. Stergiou, K. E. Psannis, B. B. Gupta, “Advanced Media-based Smart Big Data on Intelligent Cloud Systems”, IEEE Transaction on Sustainable Computing, in Press, 2018.

[34] A. A. Gnana Singh et al, "A Survey on Big Data and Cloud Computing", International Journal on Recent and Innovation Trends in Computing and Communication, vol. 7, no. 4, pp. 273-277, July 2016.

[35] O. Awodele et al, "Big Data and Cloud Computing Issues," International Journal of Computer Applications, vol. 12, no. 133, pp. 14-19, January 2016.

[36] S. Rallapallia et al, "Impact of Processing and Analyzing Healthcare Big Data on Cloud Computing Environment by Implementing Hadoop Cluster," International Conference on Computational Modeling and Security (CMS2016), pp. 16-22, December 2015.

[37] J. L. Hernandez-Ramos, M. V. Moreno, J. B. Bernabe, D. G. Carrillo, A. F. Skarmeta, “SAFIR: Secure access framework for IoT-enabled services on smart buildings”, Journal of Computer and System Sciences, vol. 81, issue: 8, pp. 1452-1463, December 2015.

[38] M. V. Moreno, L. Dufour, A. F. Skarmeta, A. J. Jara, D. Genoud, B. Ladevie, J.-J. Bezian, “Big data: the key to energy efficiency in smart buildings”, Soft Computing, vol. 20, issue: 5, pp. 1749-1762, May 2016.

[39] J. Shah, B. Mishra, “Customized IoT Enabled Wireless Sensing and Monitor-ing Platform for Smart Buildings”, Procedia Technology, vol. 23, pp. 256-263, February 2016.

[40] Hatem Ben Sta, “Quality and the efficiency of data in “Smart-Cities””, Fu-ture Generation Computer Systems, vol. 74, pp. 409-416, 2017.

[41] J. Li, Y. Zhang, X. Chen, Y. Xiang, “Secure attribute-based data sharing for resource-limited users in cloud computing”, Elsevier, Computers & Security, vol. 72, pp. 1-12, January 2018.

[42] M. Qiu, K. Gai, B. Thuraisingham, L. Tao, H. Zhao, “Proactive user-centric secure data scheme using attribute-based semantic access controls for mobile clouds in financial industry”, Elsevier, Future Generation Computer Systems, vol. 80, pp. 421-429, March 2018.

[43] M. Elhoseny, G. Ramirez-Gonzalez, O. M. Abu-Elnasr, S. A. Shawkat, Arunkumar N, A. Farouk, “Secure Medical Data Transmission Model for IoT-Based Healthcare Systems”, IEEE Access, vol. 6, pp. 20596 – 20608, March 2018.

[44] E. B. Hamida, H. Noura, W. Znaidi, “Security of Cooperative Intelligent Transport Systems: Standards, Threats Analysis and Cryptographic Countermeasures”, Electronics, vol. 4, issue: 3, pp. 380-423, July 2015.

[45] A. V. Dastjerdi, R. Buyya, “Fog Computing: Helping the Internet of Things Realize Its Potential”, IEEE, Computer, vol. 49, issue: 8, August 2016.

[46] M. Barreno, B. Nelson, R. Sears, A. D. Joseph, J. D. Tygar, “Can machine learning be secure?”, ACM, in Proceedings of he 2006 ACM Symposium on Information, computer and communica-tions security, ASIACCS '06, pp. 16-25, 21 -24 March 2006, Tai-pei, Taiwan.

[47] X. Liao, L. Ding, Y. Wang, “Secure Machine Learning, A Brief Overview”, IEEE, in Proceedings of 2011 Fifth International Conference on Secure Software Integration and Reliability Im-provement – Companion, 27-29 June 2011, Jeju Island, South Korea.

[48] M. Barreno, B. Nelson, A. D. Joseph, J. D. Tygar, “The security of machine learning”, Springer, Machine Learning, vol. 81, is-sue 2, pp. 121-148, November 2010.

[49] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, K. Seth, “Practical Se-cure Aggregation for Privacy-Preserving Machine Learning”, ACM, in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS ’17, pp. 1175-1191, 30 October – 3 November 2017, Dallas, Texas, USA.

[50] A. Demontis, M. Melis, B. Biggio, D. Maiorca, D. Arp, K. Rieck, I. Corona, G. Giacinto, F. Roli, “Yes, Machine Learning Can Be More Secure! A Case Study on Android Malware Detection”, IEEE Transactions on Dependable and Secure Computting, pp. 1-1, Early Access, May 2017.

[51] Y.-B. Sheng, L. Zhou, “Distributed secure quantum machine learning”, Else-vier, Science Bulletin, vol. 64, issue 14, pp. 1025-1029, July 2017.

[52] G. Bello-Orgaz, J. J. Jung, D. Camacho, “Social big data: Recent achieve-ments and new challenges”, Information Fusion, vol. 28, pp. 45-59, 2016.

[53] A. Gandomi, M. Haider, “Beyond the hype: Big data concepts, methods, and analytics”, International Journal of Information Management, vol. 35, no. 2, pp. 137-144, 2015.

[54] N. Kaur, S. K. Sood, “Dynamic resource allocation for big data streams based on data characteristics (5Vs)”, International Journal of Network Man-agement, vol. 27, issue 4, May 2017.

[55] H. Hu, Y. Wen, T. S. Chua, X. Li, “Toward scalable systems for big data analytics: A technology tutorial”, IEEE Access, vol. 2, pp. 652-687, 2014.

[56] V. Rajaraman, “Big data analytics.”, Resonance, vol. 21, no. 8, pp. 695-716, 2016.

[57] A. Alexandrov, R. Bergmann, S. Ewen, J. C. Freytag, F. Hues-ke, A. Heise, A., F. Naumann, “The Stratosphere platform for big data analytics”, The VLDB Journal, vol. 23, no. 6, pp. 939-964, 2014.

[58] O. Kwon, N. Lee, B. Shin, “Data quality management, data usage experience and acquisition intention of big data analyt-ics”, International Journal of Information Management, vol. 34, no. 3, pp. 387-394, 2014.

[59] K. Müller et al, "3D High-Efficiency Video Coding for Multi-View Video and Depth Data", IEEE Transactions on Image Processing, vol. 9, no. 22, pp. 3366-3378, September 2013.

[60] L. Shen et al, "An Effective CU Size Decision Method for HEVC Encoders", IEEE Transactions on Multimedia, vol. 2, no. 15, pp. 465-470, February 2013.

[61] Jens-Rainer Ohm et al, "Comparison of the Coding Efficiency of Video Coding Standards-Including High Efficiency Video Coding (HEVC)", IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 22, pp. 1669-1684, December 2012.

[62] J. M. Batalla, "Advanced multimedia service provisioning based on efficient interoperability of adaptive streaming proto-col and high efficient video coding," Journal of Real-Time Image Processing, pp. 1-12, March2015.

[63] M. Rouse, "IoT security (Internet of Things security)," IoT Agenda, 01/11/2015. [Online]. Available: http://internetofthingsagenda.techtarget.com/definition/IoT-security-Internet-of-Things-security. [Accessed 27/07/2016].

[64] A. P. Plageras, K. E. Psannis, “Algorithms for Big Data Delivery over the Internet of Things”, in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessaloniki, Greece.

[65] L. Dong, Z. Han, A. P. Petropulu, H. V. Poor, “Improving Wireless Physical Layer Security via Cooperating Relays”, IEEE Transactions on Signal Processing, VOL. 58, No. 3, March 2010.

[66] A. K. Nair et al, "Analysis of Physical layer Security via Co-operative Communication in Internet of Things," International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015), no. 24, p. 896 – 903, January 2016.

[67] W. Hu, H. Tan, P. Corke, W. C. Shih, S. Jha, “Toward trusted wireless sensor networks”, ACM Transactions on Sensor Net-works, vol. 7, issue 5, pp. 1-25, 2010.

[68] D. Huang, “Mobile cloud computing,” IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, vol. 6, no. 10, pp. 27–31, 2011.

[69] G. Md Whaiduzzaman et al, “A Study on Strategic Provision of Cloud Computing Services”, The Scientific World Journal, pp. 1-8, June 2014.

[70] S. K. Garg, S. Versteeg, R. Buyya, “A framework for ranking of cloud computing services”, Future Generation Computer Systems, vol. 29, issue: 4, pp. 1012–1023, 2013.

[71] Mohammad Haghighat et al, "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification," Expert Systems with Applications, vol. 11, no. 42, pp. 7905-7916, November 2015.

[72] Y. Kumar, R. Munjal, H.Sharma, “Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures” IJCSMS International Journal of Computer Science and Management Studies, Vol. 11, Issue 03, October 2011.

[73] R. Kaur, S. Kinger, "Analysis of Security Algorithms in Cloud Computing," International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 3, pp. 171-176, March 2014.

[74] G. Singh, S. Kinger, “Integrating AES, DES, and 3-DES Encryption Algorithms for Enhanced Data Security”, International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July 2013.

[75] A. Sachdev, M. Bhansali, "Enhancing Cloud Computing Security using AES Algorithm," International Journal of Computer Applications, vol. 9, no. 67, pp. 19-23, April 2013.

[76] P. Viswanathan, "Cloud Computing – Is it Really All That Beneficial?", abouttech, 7/7/2012. [Online]. Available: http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm. [Accessed 24/5/2017].

[77] F. Pfarr, T. Buckel, A. Winkelmann, "Cloud Computing Data Protection – A Literature Review and Analysis", in Proceedings of 47th Hawaii International Conference on System Sciences, pp. 5018-5027, 6-9 January 2014, Waikoloa, HI, USA.

[78] E. Almrot, S. Andersson, “A study of the advantages & dis-advantages of mobile cloud computing versus native environment”, Digitala Vetenskapliga Arkivet, Bachelor Thesis in Software Engineering, Blekinge Institute of Technology, Karlskrona, May 2013.

[79] S. Fremdt, R. Beck, S. Weber, "Does Cloud Computing Matter? An analysis of the Cloud Model software-as-a-service and its impact on operational agility", in Proceedings of 46th Hawaii International Conference on System Sciences 2013, pp. 1025-1034, 7-10 January 2013, Waileam Maui, Hi, USA.

[80] Blog: Follow what's happening at Get Cloud Services, "Mobile Cloud Computing – Pros and Cons," GetCloud Services, 23/12/2014. [Online]. Available: https://www.getcloudservices.com/blog/mobile-cloud-computing-pros-and-cons/. [Accessed 24/12/2017].

[81] E. Shi, Y. Niu, M. Jakobsoon, R. Chow, "Implicit Authentication through Learning User Behavior", ACM, in Proceedings of ISC'10 13th International Conference on Information Security, pp. 99-113, 25-28 October 2010, Boca Raton, FL, USA.

[82] The NIST definition of cloud computing, National Institute of Standards and Technology. [Accessed 24/07/2015].

[83] Huang D. Mobile Cloud Computing. IEEE COMSOC Multi-media Communications Technical Committee (MMTC) E-Letter, vol. 6, issue: 10, pp. 27–31, 2011.

[84] A. Botta et al, "Integration of Cloud Computing and Internet of Things: a Survey," Journal of Future Generation Computer Systems, pp. 1-54, September2015.

[85] T. Bhattasali, R. Chaki, N. Chaki,, "Secure and trusted cloud of things". In: India Conference (INDICON), 2013 Annual IEEE, pp. 1–6.

[86] Y. Simmhan, A. G. Kumbhare, B. Cao, V. Prasanna, "An analysis of security and privacy issues in smart grid software architectures on clouds", In: Cloud Computing (CLOUD), IEEE International Conference on. IEEE, pp. 582–589, 2011.

[87] N. Grozev, R. Buyya, "Inter-cloud architectures and application brokering: taxonomy and survey", Software: Practice and Experience, vol. 44, issue: 3, pp. 369–390, 2014.

[88] B. P. Rao, P. Saluia, N. Sharma, A. Mittal, S. V. Sharma, "Cloud computing for Internet of Things & sensing based applications", In: Sensing technology (ICST), 2012 Sixth International Conference on. IEEE, pp. 374–380, 2012.

[89] W. He, G. Yan, L. D. Xu, "Developing vehicular data cloud services in the iot environment", IEEE Transactions on Indus-trial Informatics, vol. 10, issue: 2, pp. 1587–1595, May 2014.

[90] C. Dobre, F. Xhafa, F., "Intelligent services for big data science", Future Generation Computer Systems, vol. 37, pp. 267–281, 2014.

[91] G. Aceto, A. Botta, W. de Donato, A. Pescap`e, "Cloud monitoring: A survey", Computer Networks, vol. 57, issue: 9, pp. 2093–2115, 2013.

Chapter 8

Efficient and Secure Big Data delivery in Cloud Computing

C. Stergiou, K. E. Psannis, "Efficient and Secure Big Data delivery in Cloud Computing", Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017. [DOI:10.1007/s11042-017-4590-4]

Reprinted by permission from **C. L. Stergiou**: **Springer Multimedia Tools and Applications Efficient and Secure Big Data delivery in Cloud Computing, C. Stergiou, K. E. Psannis**, **2017**

8.1 Introduction

Big Data (BD) is a new technology which rapidly growing in the telecommunications sectors, especially in the contemporary field of wireless telecommunications. Another technology that grows rapidly in the field of wireless telecommunications is Cloud Computing (CC). CC concerns an infrastructure where data storage and processing take place outside of the user's device. Both of them face security and privacy issues in their function. To improve them and to optimize their privacy and security issues conducted the present survey. In this paper, I survey BD and CC technology and their basic characteristics, with a focus on the security and privacy issues of both technologies. Specifically, I try to combine the functionality of the two technologies (i.e BD and CC) to examine the frequent features, and also to discover the benefits related to security issues of their integration. Concluding, I present a new method of an algorithm that can be used to improve Cloud Computing's security through the use of algorithms that can provide more privacy in the data related to Big Data technology. In the end, there is a survey about the challenges of the integration of BD and CC related to their security level.

8.2 Related Review

For the purpose of this paper we study and analyze previous studies in the field of Big Data technology and Cloud Computing and we examine existing work proposed both in the literature and on the Internet. Below presented the papers we have studied with their main objective.

To begin with, there are several works for the Cloud Computing Technology. An exploration of the possibilities solutions to offer a trustworthy CC environment presented in [15]. Then, the [16] proposes a data encryption model which protects the privacy and security of the data before they are uploaded in the cloud. The key consideration dealt in the proposal of [16] is the encryption schema in order to secure data with the aim to make it unintelligible for all. Regarding to [16], using AES for security over data is of great importance as it offers a variety of benefits such as less memory consumption and less computation time. Furthermore, a survey of the MCC was given in [17], which has the purpose to help general readers that have a survey of MCC with focus on its definition, its architecture and its applications. Also, it was presented the issues, the existing solutions and the approaches of them in [17]. Finally, [17] suggest the future research directions of the MCC technology. The [18] provides a thorough overview of MCC investigation, while highlighting the unique concerns in mobile cloud computing. Furthermore, a taxonomy count on the key problems in the area of mobile cloud computing and a discussion about the different methods that have been taken to tackle this issues were presented in [18]. At the end, there is a conclusion with a critical analysis of possibilities which have not yet fully met, and address directions for the future work. In [18] it was detailed the security issues which result because of the very nature of CC. Furthermore, the [19] presents the newly born solutions that presented in its literature in order to counter the security issues. Also, a brief view of security vulnerabilities in the MCC is highlighted. Finally, it presented the discussion on the open problems and future work directions. At the end, regarding the cloud computing technology, the [20] offers a study on cloud computing and appropriate algorithms for balancing the load. Also, it was given comparison between those algorithms on different properties of them. According the [20], the ACO is the best algorithm for balancing the load. Concluding the first part of related review, there is study about BD technology.

As already mentioned, the main purpose of this work is the operation of Big Data technology through the technology of Cloud Computing. At the beginning, the [21] presents a survey on the big data and cloud computing, with the importance to promote the research and development activities in the field of the big data and the cloud computing. At the end, the [21] presents a method for storing the data on cloud using the CloudSim package. Thereinafter, through [22] the authors gave a discussion about the Big Data using cloud computing. Also, in [22] explained how people adopt cloud as Cloud Technologies Mature. Furthermore, the authors gave an explanation of how Big Data and cloud is responding for user's request as Cloud and Big Dara a compelling combination. In [23] the authors discuss on the effects of data processing and analyzing big healthcare data on CC environment. The [23] suggest the use of the Hadoop, which is a framework that could process a large amounts of data sets on distributed environment, and also it can be deployed on CC environment to process the big healthcare data. Then, the [23] shows a research that focuses on the two key concepts, BD and CC, and some of the problems and possibilities which are inherent with the deployment of CC and BD services. Through this study is shown which security challenges is among the most prominent problem in CC and BD services. Finally, after there is a consideration about some of the problems related to BD and CC, a number of solutions were suggested by [24] towards improving the two key concepts that will go a long way in increasing the adoption rate of CC by organizations. The [25] explores the contemporary methods in the field of BD using cloud resources and how the medium-sized enterprises (MSEs) can take advantage of these technological methods. Finally, the results of [25] will benefit MSEs in identifying and exploring possible opportunities and moreover make clear the challenges in influencing BD. At the end, the [26] nominate a cloud-assisted differentially private video composition system which is count on allocated online learning, in order to handle privacy and other problems. Also, as regards as the sparsity and heterogeneity of big social media data, the authors of [26] propose an innovative "geometric differentially private" model, that could minimize the performance loss. Finally, the simulation which used in [26] displays the proposed algorithms exceed a number of existing methods and retain a delicate balance between the total reward and privacy preserving level.

Table 8.1 lists the findings and the concepts examined in each paper.

Year	Author	Problems	Solutions
2010	H. Takabi et al **[15]**	• Specific characteristics worsen security & privacy challenges of Cloud Computing.	• Examines the possibilities of offering a trustworthy Cloud Computing environment.
2011	H. T. Dinh et al **[17]**	• Detonating growth of mobile applications & resurgent of CC concept is considered advancement in mobile services.	• A survey of MCC, with focus on its definition, architecture & applications.
2012	N. Fernando et al **[18]**	• Intrinsic problems (e.g. resource scarcity, frequent disconnections) hinter the usage of mobile computing in its full scale.	• Categorizes the major issues in MCC & discusses different methods to solve these issues. • Careful examination of problems which have not yet been addressed & put forward ideas for future research.
2013	Sachdev & M. Bhansali **[16]**	• The bigger the number of cloud users the most frequent the malicious activity in the cloud. • Highly safe and persistent services needed.	• A data encryption model which protects the privacy and security of the data before they are uploaded in the cloud.
2015	M. Ali et al **[19]**	• Third-party cloud services have more deficiencies and more vulnerable to security threads. • Sharing the users' data outside the administrative control.	• Examines and shortly analyzes both internal and external security problems in the Mobile Cloud Computing.
2015	S. Bhavani et al **[20]**	• Load balancing is one of the cloud's issues. • A reduction in the response time and optimization of the resource utilization can be achieved balancing the load.	• The best algorithm for balancing the load is Ant Colony Optimization.
2015	S. Sathya & R. Avinash **[22]**	• How people adopt cloud as Cloud Technologies Mature.	• An explanation of how BD and cloud responds for user's demand as a compelling combination.
2015	S. Rallapalli et al **[23]**	• The healthcare organizations face the critical challenge to analyze big data. • Large amounts of data cannot be processed through conventional systems.	• Hadoop: An application which could prepare huge amounts of data in distributed environment could be deployed on cloud environment to prepare the big amount of healthcare data.
2016	O. Awodele et al **[24]**	• Security challenges are the most serious in cloud & big data services. • Issues of service level agreement.	• Shipping disk drives to cloud computing. • Use of Data mining techniques. • Use of Access control techniques.
2016	N. R. Vajjhala & E. Ramollari **[25]**	• Contemporary methods in the field of big data using cloud resources. • How the SMEs can take advantage of these technological trends.	• Cloud computing offers an alternative to SMEs shifting the burden of providing and maintaining expensive infrastructure to cloud service providers.
2016	P. Zhou et al **[26]**	• The increased usage of social media has created a new period, that of the big data. • Privacy of users' contexts & video service sellers' repositories, that are remarkably sensitive & of important commercial value.	• An innovative "geometric differentially private" scheme, that could minimize the performance loss.
2016	A. A. Gnana Singh et al **[21]**	• Promote the research and development activities in the field of the big data and the cloud computing.	• A method for storing the data on cloud using the cloudsim package.

Table 8.1: Mapping problems against referenced solutions.

8.3 Big Data

Due to its scale is much larger, big data could be defined that is a more complicated world. Big data sets advanced analytic techniques in which operate on, that called BD Analytics. Therefore, BD analytics is divided in two things, BD and analytics, in addition how those two have teamed up in order to establish one of the most profound methods in business intelligence (BI) today. The best way in order to discover better use of data through modern application (e.g. identifies the best suppliers, understand sales seasonality) [27].

8.3.1. Big Data Characteristics

The amount of data in storage usually characterizes the most definitions of big data. Instead of size there are also other important attributes that matters in big data such as data variety and data velocity. The three Vs of big data (figure 8.1), which are volume, variety, and velocity, make up an inclusive definition. More specifically, each one of the three Vs has its own specializations for analytics [27] [28].

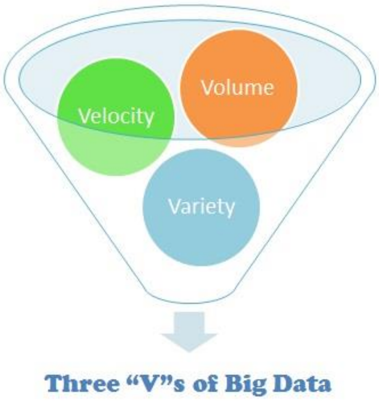


Figure 8.1: The Three Vs of Big Data.

Big Data Volume

The amount of data that is available for an organization and additionally it is not necessary for the organization to own all of the data that can access it, could be defined as the Data Volume. Regard that data volume goes up; the value of different data records would reduce in symmetry to age, type, richness, and quantity between other factors [27] [28].

Big Data Velocity

Various attributes such as the speed of data creation, agglomeration and streaming are related to the term of Data Velocity. The e-Commerce has fast increase the speed and richness of data calculated for various business transactions. The management of Data velocity is much more than a bandwidth problem; it is additionally an ingest problem (extract-transform-load) [27] [28].

Big Data Variety

The plenty of the data depiction, like text, audio, video, etc., consist the Data Variety. Important challenges which could be lead to an analytic sprawl are represented by non-aligned data structures, irreconcilable data formats and inconsequent data semantics [27] [28].

8.3.2. Applications of Big Data

In recent years, the number of mobile phone subscriptions outreaches the 4 billions all over the world, and about 2 billion people have access to the internet [29]. Between the past twenty years, about a 1 billion people throughout the world became a member of the middle class, and as an outgrowth more people are considered as literate, leading to information progress [30] [31].

Predictable & Efficient

Applications of BD importantly go up the quantity of real-time and workload-intensive transactions through the huge amounts of different data transferred. The supporting network which connects the hyperscale server architectures, insisting of thousands of nodes which in turn include various processors, might be enough to make certain this data can move rapidly and efficiently [32-34].

Holistic Network

Regarding the optimization of network performance, it is important which it takes place through the BD domain considered in the connection with the more conventional undertaking infrastructure. A holistic network approach provides some advantages as the Multitenancy, the ability to reduce duplicate costs and to leverage network staffing expertise, and finally the ease of network provisioning [32-34].

Network Partitioning

Without additional cost and complexity the division which was enabled by logical separation. Furthermore, different tasks may also need to be isolated by the use of hard partitioning on the Ethernet switch. Regarding this the tasks are totally divided at the level of data plane. For instance, the separation of the data plane could be important in order to comply with regulations and privacy needs [32-34].

Scale Out

The ability of “junior science projects” (projects of BD which may begin small) to “Scale Out” could guarantee a seamless devolution as projects grow in size and number. In addition to this, an important issue to the same degree is that network performance and ease of management remain constant as the cluster scales [32-34].

Unified Ethernet Fabrics

By influencing various paths into the network, and constantly determining the most efficient route, Unified Ethernet Fabrics empowers full link utilization. A perfect scalability could be provided by the Unified Ethernet Fabrics, since the virtual chassis architectures offer access to various switches and, simultaneously, manage them as a unique device. Moreover, by this design might be offered a predictable any-to-any latency and bandwidth for traffic between servers through the BD cluster [32-34].

8.3.3. Big Data Security & Privacy

New challenges and standards developed and created in data security issues through the development and the use of BD technology. This creates a growing need for further research on security technologies in order to be able to handle the huge amount of data and to ensure effective. Technologies in order to secure data are slow when applied to large amounts of data.

Regarding the Table 8.2 we could conclude that even the most efficient algorithms give an encryption rate of 64.3MB/s [32-34]. Thus, in the sector of BD technology, in which the need of large amounts of data needs to be transferred we could observe an important bottle neck for encryption like huge amounts of data. This is harmful to the nature of BD that have real time processing and outcomes.

The flowcharts of the algorithms which have been studied in this paper are presented in Table 8.3 and Table 8.4 [32-34].

Algorithm	*Key length*	*Megabytes processed*	*Block size*	*Rounds*	*Time taken*	*MB per second*
Blowfish	32-448 bits	256	64 bits	16	3,976	64,386
DES	56 bits	128	64 bits	16	5,998	21,340
3-DES	56, 112 or 168 bits	128	64 bits	48	6,159	20,783
AES	128, 192 or 256 bits	256	128 bits	10, 12 or 14	4,196	61,010
RSA	1025 – 4096 bits	300	512 bits	1	1175,783	10,900

Table 8.2: Encryption Rates of popular Algorithms.

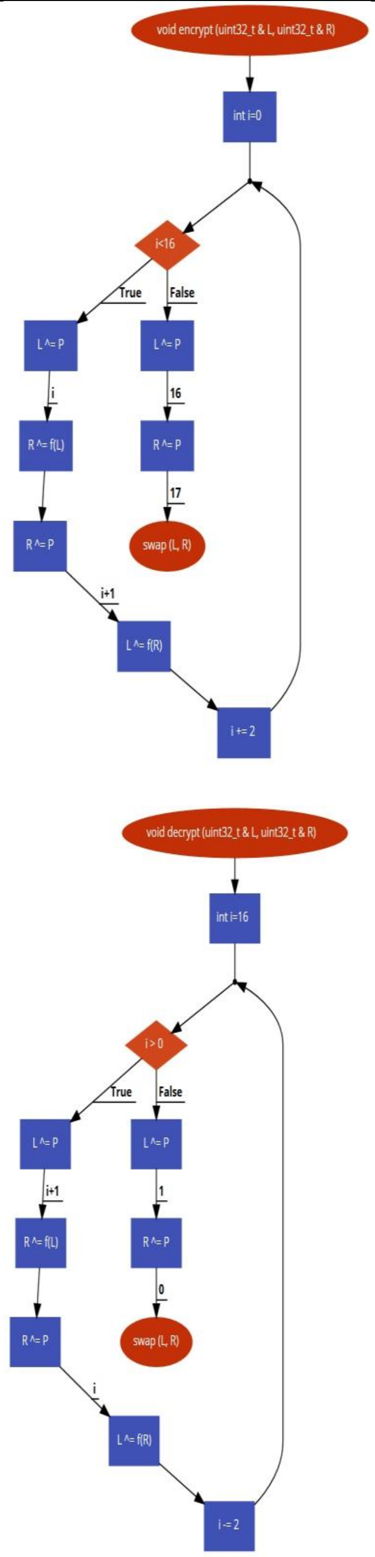
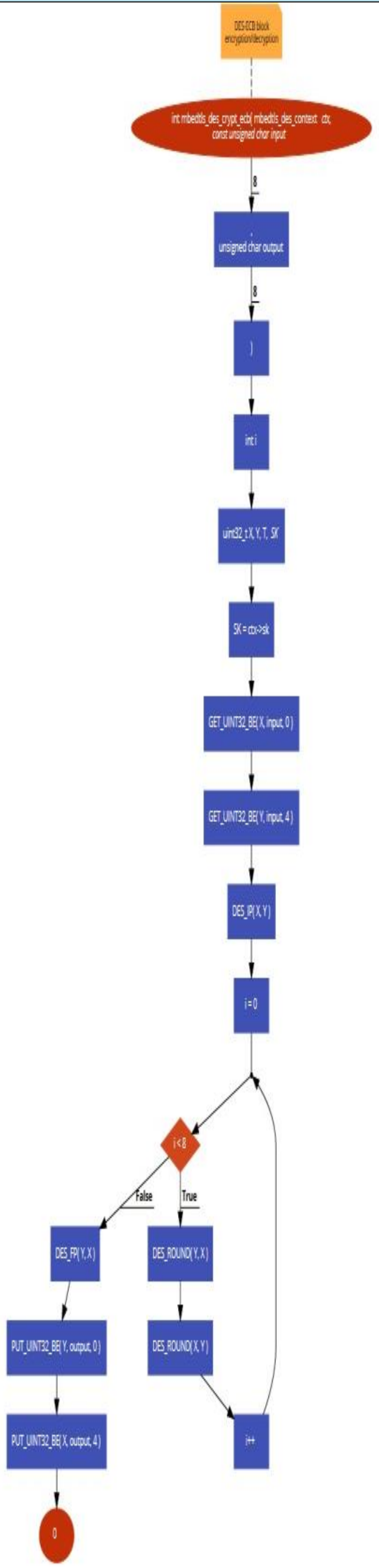
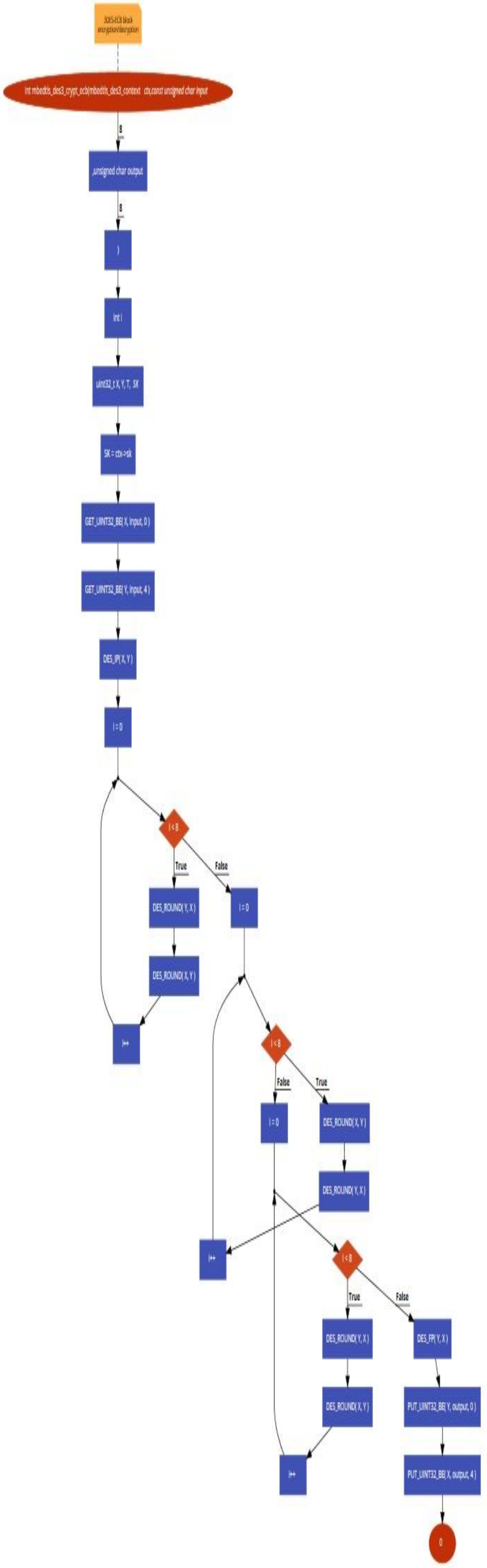
Blowfish Algorithm *(Encryption-Decryption only)*	***DES Algorithm*** *(Encryption-Decryption only)*	***3-DES Algorithm*** *(Encryption-Decryption only)*
 void encrypt (uint32_t & L, uint32_t & R) int i=0 i<16 True False L ^= P L ^= P i 16 R ^= f(L) R ^= P 17 R ^= P swap (L, R) i+1 L ^= f(R) i += 2 void decrypt (uint32_t & L, uint32_t & R) int i=16 i > 0 True False L ^= P L ^= P i+1 1 R ^= f(L) R ^= P 0 R ^= P swap (L, R) i L ^= f(R) i -= 2	 DES-ECB block encryption/decryption int mbedtls_des_crypt_ecb(mbedtls_des_context ctx, const unsigned char input 8 unsigned char output 8) int i uint32_t X, Y, T, SK SK = ctx->sk GET_UINT32_BE(X, input, 0) GET_UINT32_BE(Y, input, 4) DES_IP(X, Y) i = 0 i < 8 False True DES_FP(Y, X) DES_ROUND(Y, X) PUT_UINT32_BE(Y, output, 0) DES_ROUND(X, Y) PUT_UINT32_BE(X, output, 4) i++ 0	 3DES-ECB block encryption/decryption int mbedtls_des3_crypt_ecb(mbedtls_des3_context ctx,const unsigned char input 8 ,unsigned char output 8) int i uint32_t X, Y, T, SK SK = ctx->sk GET_UINT32_BE(X, input, 0) GET_UINT32_BE(Y, input, 4) DES_IP(X, Y) i = 0 i < 8 True False DES_ROUND(Y, X) i = 0 DES_ROUND(X, Y) i++ i < 8 False True i = 0 DES_ROUND(X, Y) DES_ROUND(Y, X) i++ i < 8 True False DES_ROUND(Y, X) DES_FP(Y, X) DES_ROUND(X, Y) PUT_UINT32_BE(Y, output, 0) i++ PUT_UINT32_BE(X, output, 4) 0

Table 8.3: Blowfish, DES & 3DES Algorithms (part of their code).

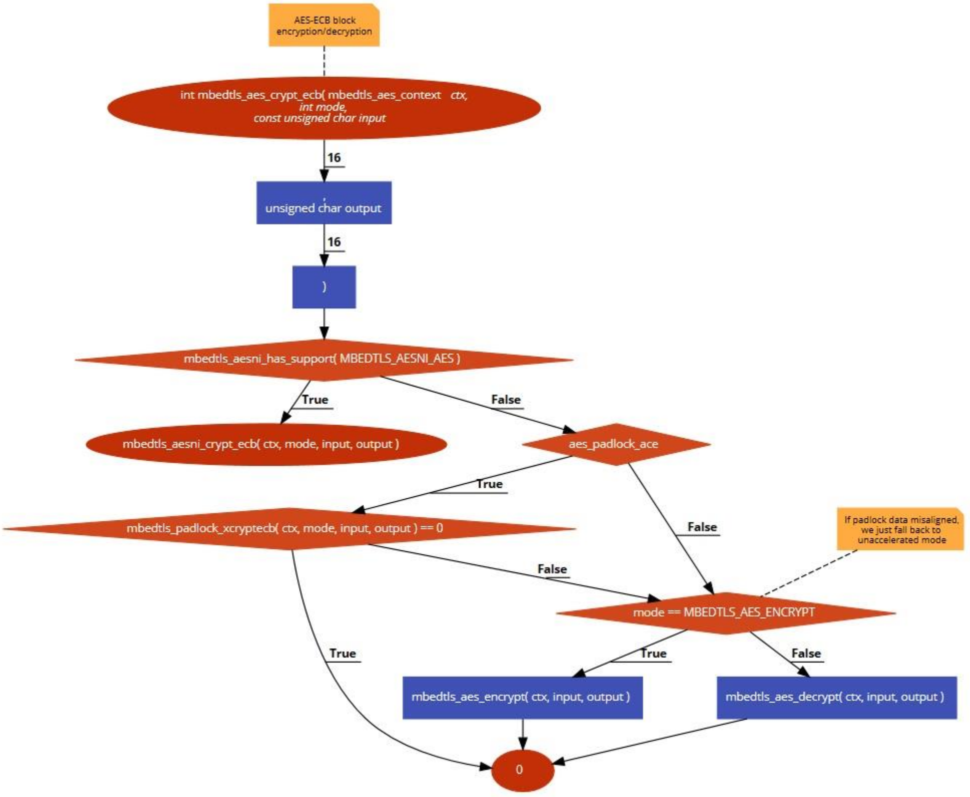
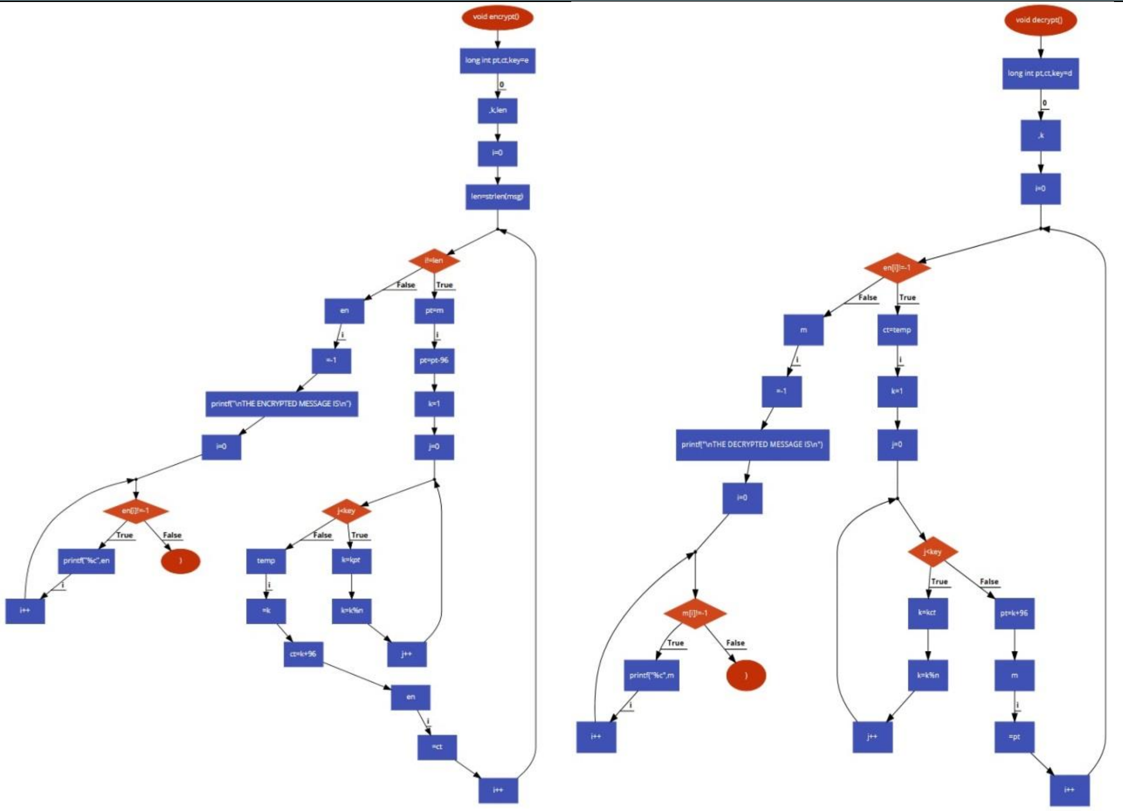
***AES Algorithm** (Encryption-Decryption only)*
 AES-ECB block encryption/decryption int mbedtls_aes_crypt_ecb(mbedtls_aes_context *ctx*, *int mode*, *const unsigned char input* 16 unsigned char output 16) mbedtls_aesni_has_support(MBEDTLS_AESNI_AES) True False mbedtls_aesni_crypt_ecb(ctx, mode, input, output) aes_padlock_ace True False mbedtls_padlock_xcryptecb(ctx, mode, input, output) == 0 If padlock data misaligned, we just fall back to unaccelerated mode False mode == MBEDTLS_AES_ENCRYPT True True False mbedtls_aes_encrypt(ctx, input, output) mbedtls_aes_decrypt(ctx, input, output) 0
***RSA Algorithm** (Encryption-Decryption only)*
 void encrypt() long int pt,ct,key=e 0 ,k,len i=0 len=strlen(msg) i!=len False True en i =-1 printf("\nTHE ENCRYPTED MESSAGE IS\n") i=0 en[i]!=-1 True False printf("%c",en i i++) pt=m i pt=pt-96 k=1 j=0 j<key False True temp i =k ct=k+96 k=kpt k=k%n j++ en i =ct i++ void decrypt() long int pt,ct,key=d 0 ,k i=0 en[i]!=-1 False True m i =-1 printf("\nTHE DECRYPTED MESSAGE IS\n") i=0 m[i]!=-1 True False printf("%c",m i i++) ct=temp i k=1 j=0 j<key True False k=kct k=k%n j++ pt=k+96 m i =pt i++

Table 8.4: AES & RSA Algorithms (part of their code).

8.4 Cloud Computing

Computing, storage, services, and applications over the Internet is provided by CC. The integration of CC technology with mobile devices in order to make the mobile devices resource-full in terms of computational power, memory, storage, energy, and context awareness, is defined as MCC (figure 8.2). This technology is the result of interdisciplinary methods integrating MC with CC. So, this transdisciplinary domain is additionally mentioned as MCC [35] [36].

Figure 8.2: Cloud Computing Technology.

8.4.1. Cloud Computing Features

As all technologies, so the CC technology has some features which determine its function. These characteristics are analyzed and outlined consequently.

<u>Storage over Internet</u>

The technology framework which uses Transmission Control Protocol/Internet Protocol (TCP/IP) networks in order to link servers and storage devices and with the aim to accommodate storage solution deployment could be defined as Storage over Internet. This technology is also publicly recognized as Storage over Internet Protocol (SoIP) [33] [34] [37].

<u>Service over Internet</u>

The most important aim of the Service over Internet is to be devoted to facilitate customers throughout the world with the aim to transform aspirations through achievements by harnessing the Internet’s efficiency, speed and ubiquity [33] [34].

<u>Applications over Internet</u>

The programs which produced to make useful operations of a present manual task, or virtually anything, and that perform their operation on the cloud server through an internet connection as well as the conventional model of a program which need to be installed and operate on a local computer are the Cloud Applications, or as scientific known, Applications over Internet [32-34].

Energy Efficiency

The path of managing and restraining the increase in energy consumption could be defined as Energy Efficiency. For instance, when a Compact Florescent Light (CFL) bulb uses less energy (1/3 to 1/5) than an incandescent bulb to generate the same quantity of lights, the Compact Florescent Light (CFL) is considered to be more energy efficient [32-34].

Computationally Capable

The technology of the MCC enables all the services of computational clouds that are influencing the computationally strenuous and omnipresent mobile applications. So, a system is admitted as computationally capable as long as it achieves the needs, to offer us the results we want, by progressing the correct calculations [33] [34].

8.4.2. Disadvantages of Cloud Computing

CC has a number of trade offs which need to be minimized over the years with the aim to reach a better and more ideal use.

Security

Security is one major issue of the MCC technology. In MCC technology must be considered that all the sensitive information could be surrendered to various third-party cloud service providers. Therefore, regarding to this when someone need to rely information to a cloud provider must be absolute sure that the provider will keep the information totally safe [38] [39].

Connectivity

A critical and vital issue of MCC is Internet connection. The user must be certain with the choice of a good and reliable internet provider [40].

Performance

The performance of the MCC is another primary issue. A number of users need to feel that the performance of services is good enough as in native applications [41] [42].

Latency (Delay)

Latency, or as also known Delay, is defined as the time involved in offloading the computation and getting back the results from the adjacent infrastructure or cloud, in the field of MCC.

Privacy

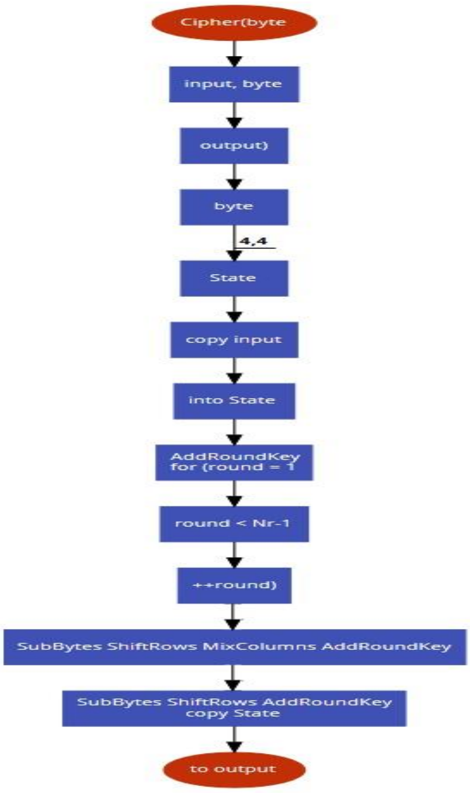
When a user adopts MCC a major bottleneck which could be considered is privacy. Hence, in order to gain consumers reliance in the MC, the application models may support application progress with privacy protection, and unrestricted authentication mechanisms [39] [43].

8.4.3. Cloud Computing Security Model

In order to offer secure communication through the network, encryption algorithm plays an important role. It is a valuable and fundamental mechanism for the protection of the data. Encryption algorithm converts the data into scrambled form with the use of “a key” and only the user have the key to decrypt the data. Regarding the researches that have been made, an important encryption technique is the Symmetric key Encryption. In this key encryption method, in order to encrypt and decrypt the data only one key is used. In this encryption technique the most used algorithm is the AES [44] [45] [46].

In order to replace the DES algorithm NIST recommended a new encryption standard which is the AES (Advanced Encryption Standard) algorithm. The AES algorithm block ciphers. AES used a number of keys in order to encrypt and decrypt, with length of 128, 192, or 256 bits, but as a default the key length is 256 bits. Depending to the key size the AES algorithm encrypts the data blocks of 128 bits in 10, 12 and 14 rounds. Additionally, AES has been tested with attention for a huge number of security applications [9] [47] [48].

AES algorithm considered as better than others for a number of reasons, which is follows [16] [47]:


Cipher(byte
input, byte
output)
byte
4,4
State
copy input
into State
AddRoundKey for (round = 1
round < Nr-1
++round)
SubBytes ShiftRows MixColumns AddRoundKey
SubBytes ShiftRows AddRoundKey copy State
to output

- ✓ A very good software performance resulted by AES’s innate parallelism facilitates efficient use of processor resources.
- ✓ Fast key set up time & good key agility provided by AES.
- ✓ AES is suitable for limited space environments instead of the less memory requirements for implementation.
- ✓ AES is protected by the no serious weak keys.
- ✓ Key and block sizes which are greater than 128 bits are supported by the AES.
- ✓ AES is currently protected by linear and differential cryptanalysis attacks.

Additionally, there is an important encryption technique from the Asymmetric key Encryption. Instead of Symmetric key encryption, two keys, one private and one public, is used in the Asymmetric key encryption. The private key is used for the decryption and the public key is used for encryption [44] [45].

8.5 Big Data in Cloud Computing

In order to be able for every user to manage and process big amounts of data everywhere and every time a new challenge created. This challenge is to use BD in CC. Regarding to the related research and by surveying the two technologies we reached that through this integration there are new challenges generated.

Cloud Computing	***Storage over Internet***	***Service over Internet***	***Applications over Internet***	***Energy Efficiency***	***Computationally Capable***
Volume		**X**	**X**	**X**	
Velocity	**X**		**X**	**X**	**X**
Variety		**X**	**X**		**X**

Table 8.5: Contributions of Big Data in Cloud Computing.

Table 8.5 exhibits the key characteristic of the two technologies (CC and BD) which have been studied and used in order to use them for the experimental proposal. Based on the study conducted, the key characteristic of BD technology which contributes more with the characteristics of CC technology is Velocity. Velocity contributes four from the five key characteristics of CC. Also, another thing that we can observe from Table 8.5 is that the characteristic Applications over Internet contributed from all the key characteristics of BD.

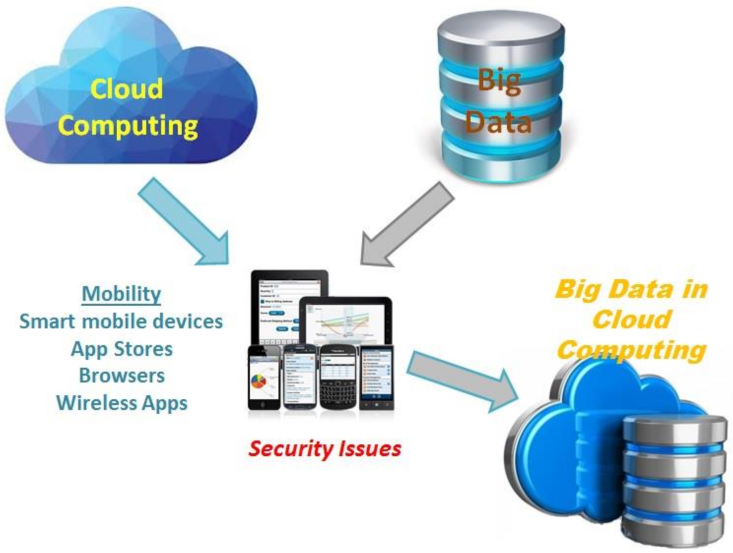


Figure 8.3: Big Data in Cloud Computing

Thereafter, an important issue in the integration BD in CC is Security of data stored in the cloud (figure 8.3). A large number of security issues come through the CC technology. This would be as a result of the fact that it encompasses a number of technologies which might contain networks, operating systems, etc. Therefore, security problems of these systems and technologies subsist in CC. The security problems related to CC devices and environments.

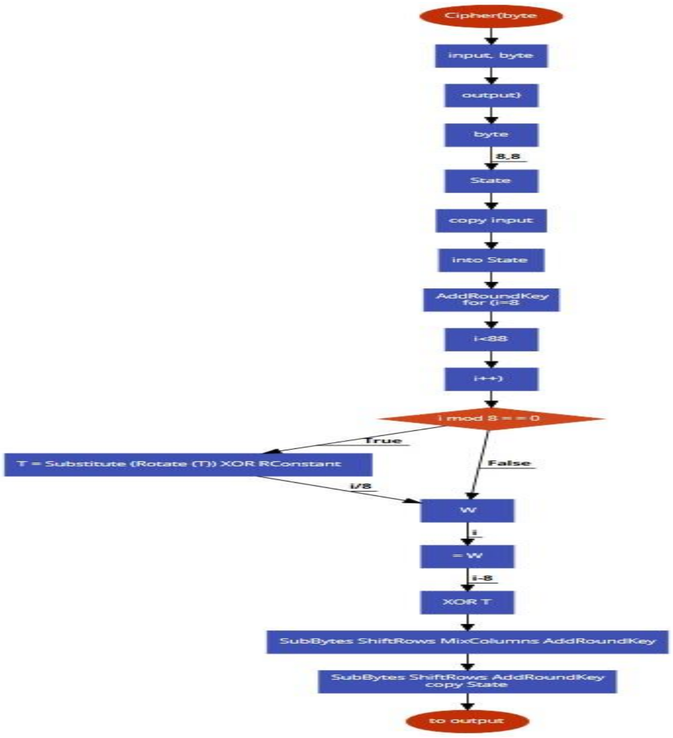
Regarding the BD security issues, they are overblown by the three key features of BD that are the three Vs (volume, variety, velocity).

Big Data Applications	**Security**	**Connectivity**	**Performance**	**Latency (Delay)**	**Privacy**
Predictable & Efficient	X		X	X	X
Holistic Network	X	X	X		X
Network Partitioning	X		X	X	X
Scale Out	X	X			X
Unified Ethernet Fabrics	X	X	X		X

Table 8.6: Contribution of Big Data Applications by Cloud Computing trade offs.

Table 8.6 lists the BD Applications and how they contributed by the CC trade offs. Affirming the earlier study, we can observe that Security and Privacy are the CC's trade offs which contribute more the BD Applications. In contrast, Latency is the CC's trade off which contributes less the BD Applications. Based on these conclusions, we can confirm that we must propose a new model which strives to improve the issues of Security and Privacy.

Relying on this study in encryption algorithms of two technologies we propose a new part of flowchart model related to the original AES flowchart. This model of AES flowchart uses the original key consists of 256 bits/16 bytes which are demonstrated as a matrix of 8x8.



Considering the benefits of the security models and algorithms of BD and CC technologies we can observe that we can have a beneficial use of integration those two technologies.

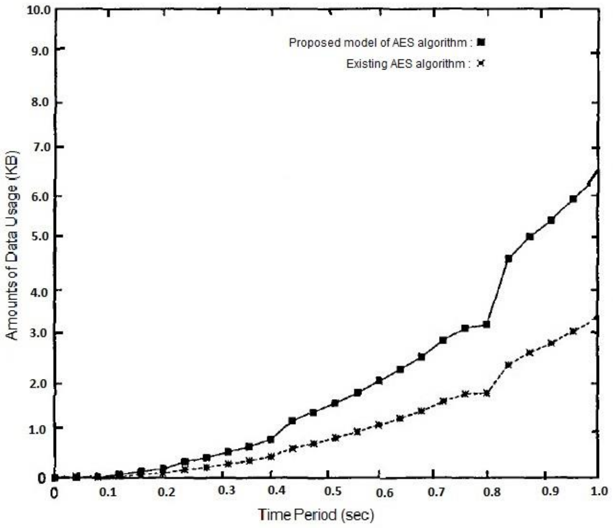


Figure 8.4: Security level of encryption algorithms of measurement used for the study of AES model algorithm.

Figure 8.4 shows, the measurements that have been through time. As we can observe by this figure the more often is the combined use of the algorithms, the higher level of security of the data usage we get every time. The upper line represents our proposed flowchart-model of AES algorithm and the other (down line) represents the existing AES algorithm.

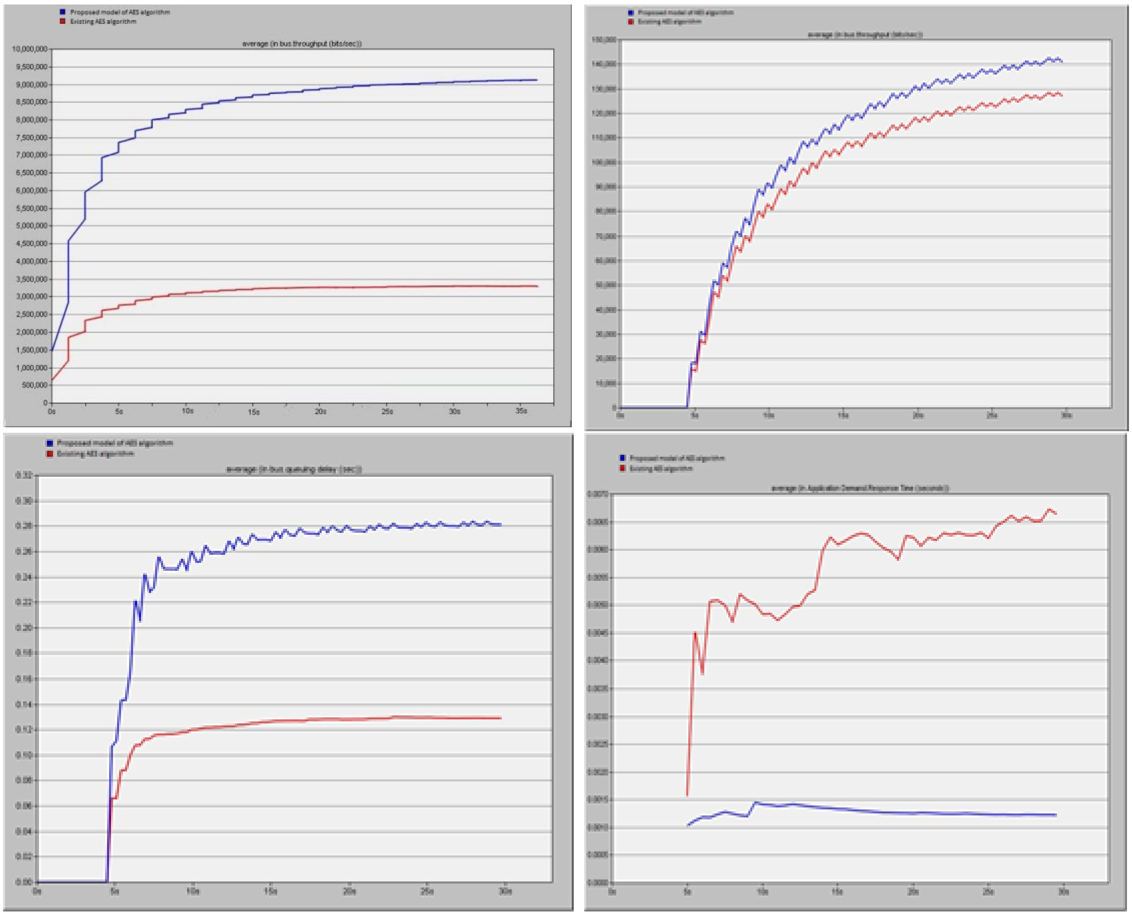


Figure 8.5: Security level of encryption algorithms of measurement used for the study of AES model algorithm (bits per seconds).

Figure 8.5 demonstrates the measurements that have been through time regarding the different amount of bits used. As we can observe by those figures represented in Figure 8.5, the more often is the combined use of the algorithms, the higher level of security of the data usage we get every time. The blue line represents our proposed model of AES algorithm and the red represents the existing AES algorithm.

8.6 Chapter Summary

The CC technology provides many possibilities, but in addition to this places quite a lot of restrictions as well. This technology mentions to an infrastructure where both the data storage and processing occur outside of the user’s device. In this paper, we survey BD and CC technology and their basic characteristics, with a focus on the security and privacy issues of both technologies. Moreover, we have tried to combine the functionality of the two aforementioned technologies (i.e BD & CC) with the aim to examine the frequent characteristics, and moreover to discover the benefits related in security issues of their integration.

The main goal of this paper is to try to combine the functionality of the BD and CC technologies with the aim to examine the frequent characteristics, and also to discover the benefits related in security issues of their integration. This could be take place by the presentation of a new method of an algorithm that can be used for the purpose of improving CC's security through the use of algorithms that can provide more privacy in the data related to BD technology. Moreover, we survey the security challenges of the integration of those technologies. This can be the field of future research on the integration of those two technologies, and why not to have a huge improvement of their security and privacy issues in order to have a better use of them.

8.7 Chapter References

[9] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, "Secure integration of IoT and Cloud Computing", Elsevier, Future Generation Computer Systems, vol. 78, part 3, pp. 964-975, January 2018. [DOI:10.1016/j.future.2016.11.031]

[15] H. Takabi, J. B. D. Joshi, "Security and Privacy Challenges in Cloud Computing Environments", IEEE Security & Privacy, vol. 8, Issue: 6, pp. 24-31, November-December 2010. [DOI: 0.1109/MSP.2010.186]

[16] A. Sachdev, M. Bhansali, "Enhancing Cloud Computing Security using AES Algorithm", International Journal of Computer Applications, vol. 9, no. 67, pp. 19-23, April 2013.

[17] H. T. Dinh, C. Lee, D. Niyato, P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches", Wiley, Wireless Communications & Mobile Computing, vol. 13, Issue: 18, pp. 1587-1611, October 2011. [DOI: 10.1002/wcm.1203]

[18] N. Fernando, S. W. Loke, W. Rahayu, "Mobile cloud computing: A survey", Elsevier, Future Generation Computer Systems, vol. 29, Issue: 1, pp. 84-106, January 2013. [DOI: 10.1016/j.future.2012.05.023]

[19] M. Ali, S. U. Khan, A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges", Elsevier, Information Sciences, vol. 305, pp. 357-383, June 2015. [DOI: 10.1016/j.ins.2015.01.025]

[20] Prof. S. Bhavani, A. Hatwal, U. Mittal, ""Study on Cloud Computing and Different Load Balancing Algorithms in Cloud Computing", International Journal of Emerging Research in Management &Technology, vol. 4, Issue: 5, pp. 331-336, May 2015. [ISSN: 2278-9359]

[21] R. Iqbal, F. Doctor, B. More, "Big Data analytics: Computational intelligence techniques and application areas", International Journal of Information Management, pp. 1-11, June 2016. [DOI: 10.1016/j.ijinfomgt.2016.05.020]

[22] S. Sathya, R. Avinash, “Big Data and Cloud Computing”, in Proceedings of Rathinam College National Conference, 2015, Pollachi Road, Echanari, India.

[23] S. Rallapallia, R. R. Gondkar, U. P. K. Ketavarapu, “Impact of Processing and Analyzing Healthcare Big Data on Cloud Computing Environment by Implementing Hadoop Cluster”, Elsevier, Procedia Computer Science, vol. 85, pp. 16-22, 2016. [DOI: 10.1016/j.procs.2016.05.171]

[24] A. Oludele, A. Izang, S. O Kuyoro, F. Y. Osisanwo, “Big Data and Cloud Computing Issues”, International Journal of Computer Applications, vol. 12, no. 133, pp. 14-19, January 2016. [DOI: 10.5120/ijca2016907861]

[25] Dr. N. R. Vajjhala, Dr. E. Ramollari, “Big Data using Cloud Computing - Opportunities for Small and Medium-sized Enterprises”, European Journal of Economics and Business Studies, vol. 2, issue: 1, pp. 129-137, January-April 2016. [ISSN 2411-4073]

[26] P. Zhou, Y. Zhou, D. Wu, H. Jin, “Differentially Private Online Learning for Cloud-Based Video Recommendation With Multimedia Big Data in Social Networks”, IEEE Transactions on Multimedia, vol. 18, Issue: 6, pp. 1217-1229, June 2016. [DOI: 10.1109/TMM.2016.2537216]

[27] Cloud News Daily, "Guide to Big Data Analytics: Platforms, Software, Companies Tools, Solutions and Hadoop," Cloud News Daily, 12/12/2015. [Online]. Available: http://cloudnewsdaily.com/big-data-analytics/. [Accessed 21/5/2016].

[28] P. Russom, “Big Data Analytics”, TDWI RESEARCH, TDWI Best Practices Report, Fourth Quarter 2011, USA.

[29] “Data, data everywhere”, The Economist. [Online], Posted: 27/02/2010, Available: https://www.economist.com/special-report/2010/02/27/data-data-everywhere [Accessed 26/09/2015]

[30] M. Hilbert, P. López, “The World’s Technological Capacity to Store, Communicate, and Compute Information”, AAAS, Science, vol. 332, issue: 6025, pp. 60–65, April 2011. [DOI:10.1126/science.1200970]

[31] O. Badve, B. B. Gupta BB, S. Gupta, “Reviewing the security features in contemporary security policies and models for multiple platforms”, IGI Global, Book Chapter, Handbook of research on Modern Cryptographic Solutions for Computer and Cyber Security, pages: 26, USA, 2016. [DOI: DOI: 10.4018/978-1-5225-0105-3.ch020]

[32] L. Borovick, R. L. Villars, "The Critical Role of the Network in Big Data Applications", IDC Analyze the Future, White Paper, Sponsored by: Cisco Systems pp. 1-12, April 2012.

[33] Md Whaiduzzaman, M. N. Haque, Md R. K. Chowdhury, A. Gani "A Study on Strategic Provision of Cloud Computing Services", The Scientific World Journal, pp. 1-8, June 2014.

[34] S. K. Garg, S. Versteeg, R. Buyya, "A framework for ranking of cloud computing services", Elsevier, Future Generation Computer Systems, vol. 29, issue: 4, pp. 1012–1023, June 2013. [DOI: 10.1016/j.future.2012.06.006]

[35] D. Huang, "Mobile cloud computing," IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, vol. 6, no. 10, pp. 27–31, 2011.

[36] M. A. Alsmirat, Y. Jararweh, I. Obidat, B. B. Gupta, "Internet of Surveillance: A Cloud supported Large Scale Wireless Surveillance System", Springer, The Journal of Supercomputing, vol. 73, pp. 973-992, September 2016. [DOI: 10.1007/s11227-016-1857-x]

[37] G. Skourletopoulos, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, J. N. Sahalos, "An Evaluation of Cloud-Based Mobile Services with Limited Capacity: A Linear Approach", Springer, Soft Computing, vol. 21, issue: 16, pp. 4523-4530, August 2017. [DOI: 10.1007/s00500-016-2083-4]

[38] P. Viswanathan, "Cloud Computing – Is it Really All That Beneficial?", abouttech, 7/7/2012. [Online]. Available: http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm. [Accessed 24/5/2017].

[39] F. Pfarr, T. Buckel, A. Winkelmann, "Cloud Computing Data Protection – A Literature Review and Analysis", in Proceedings of 47th Hawaii International Conference on System Sciences, pp. 5018-5027, 6-9 January 2014, Waikoloa, HI, USA.

[40] E. Almort, S. Andersson, "A study of the advantages & disadvantages of mobile cloud computing versus native environment", Bachelor Thesis in Software Engineering, School of Computing, Blekinge Institute of Technology, Sweden, May 2013.

[41] S. Fremdt, R. Beck, S. Weber, "Does Cloud Computing Matter? An analysis of the Cloud Model software-as-a-service and its impact on operational agility", in Proceedings of 46th Hawaii International Conference on System Sciences, pp. 1025-1034, 7-10 January 2013, Wailea, Maui, HI, USA. [DOI: 10.1109/HICSS.2013.182]

[42] Blog: Follow what's happening at Get Cloud Services, "Mobile Cloud Computing – Pros and Cons", GetCloud Services, 23/12/2014. [Online]. Available: https://www.getcloudservices.com/blog/mobile-cloud-computing-pros-and-cons/. [Accessed 24/12/2017].

[43] E. Shi, Y. Niu, M. Jakobsson, and R. Chow, "Implicit authentication through learning user behavior", Springer, in Proceedings of International Conference on Information Security (ISC 2010), Information Security pp 99-11, 25-28 October 2010, Boca Raton, FL, USA. [DOI: 10.1007/978-3-642-18178-8_9]

[44] Y. Kumar, R. Munjal, H. Sharma, "Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures", IJCSMS International Journal of Computer Science and Management Studies, Vol. 11, Issue: 3, October 2011.

[45] R. Kaur, S. Kinger, "Analysis of Security Algorithms in Cloud Computing", International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 3, pp. 171-176, March 2014.

[46] S. Veluru, Y. Rahulamathavan, B. B. Gupta, M. Rajarajan, "Privacy Preserving Text Analytics: Research Challenges and Strategies in Name Analysis," Book on Securing Cloud-Based Databases with Biometric Applications, IGI-Global's Advances in Information Security, Privacy, and Ethics (AISPE) series, 2014.

[47] D. S. Abdul Elminaam, H. M. Abdul Kader, M. M. Hadhoud, "Performance Evaluation of Symmetric Encryption Algorithms", IJCSNS International Journal of Computer Science and Network Security, vol. 8, No. 12, pp. 280-286, December 2008.

[48] G. Singh, S. Kinger, "Integrating AES, DES, and 3-DES Encryption Algorithms for Enhanced Data Security", International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July 2013.

Chapter 9

Algorithms for efficient digital media transmission over IoT and cloud networking

C. Stergiou, K. E. Psannis, A. P. Plageras, Y. Ishibashi, B.-G. Kim, "Algorithms for efficient digital media transmission over IoT and cloud networking", KoreaScience, Journal of Multimedia Information System, vol. 5, issue: 1, pp. 27-34, March 2018. [DOI: 10.9717/JMIS.2018.5.1.27]

9.1 Introduction

In recent years, with the blooming of the Internet of Things (IoT) and Cloud Computing (CC), researchers have begun to discover new methods of technological support in all areas (e.g. health, transport, education, etc.). In this paper, to achieve a type of network that will provide more intelligent media-data transfer new technologies were studied. Additionally, we have been studied the use of various open-source tools, such as CC analyzers and simulators. These tools are useful for studying the collection, storage, management, processing, and analysis of large volumes of data. The simulation platform which has been used for our research is CloudSim, which runs on Eclipse software. Thus, after measuring the network performance with CloudSim, we also use the Cooja emulator of the Contiki OS, intending to confirm and access more metrics and options. More specifically, we have implemented a network topology from a small section of the script of CloudSim with Cooja, so that we can test a single network segment. The results of our experimental procedure show that there are not duplicated packets received during the procedure. This research could be a start point for better and more efficient media data transmission.

9.2 Related Work

In this section we present related works to our research. By studying the areas of collection, delivery, management, and analysis of large-scale data (Big Data), it is concluded that data centers are responsible for everyone since everything that happens to them will affect us all. So, in [14] is presented, through several open source platforms (e.g. Arduino), the implemented data center environmental monitoring system. The system's architecture design is the implementation key of success. With the implemented design and through the Internet we can identify in real-time the system logs and status. As an extension of that system is proposed the monitoring of real-time Big Data through HTML5 charts.

A region based approach is presented in [15], where Jun-Ho Huh and Kyungryong Seo discuss about efficient power consumption through the technologies and techniques of Smart Grid. The main focus area of this research is the Programmable Logic Controller (PLC) technology in conjunction with power lines for the transmission of data in a network since it is an efficient and low-cost solution for efficient metering. The results from the analysis of the implemented PLC-based power-aware home network system design, using OPNET Modeler 14.5 PL8, were analyzed and compared to those of IEEE 802.11 WLAN MAC.

In another region based approach [16], a novel power-aware routing protocol is proposed. With this protocol and a mechanism which controls the delays, researchers maximize the lifetime of every node in an Ad-hoc network system. NS-2 was the simulator used for the verification of the network.

As is known, with the blooming of Big Data, the Cloud Computing (CC) also blossomed. However, there are open issues and challenges in this technology, for some of which are provided solutions by several researchers. In [17] there is an attempt to solve the problems of ignoring the content of multimedia and the difficulty in implementing solutions for the cloud platform. So, researchers proposed a new distributed multimedia programming model for its implementation on different service platforms and different multimedia applications. Also, an algorithm for decision making by users, based on local information, is also proposed.

One of the most challenging fields of Multi-clouds is the efficient workflow scheduling. So, in [18] researchers proposed an algorithm (Multi-Clouds Partial Critical Paths, MCPCPP) for Big data scheduling in Multi-clouds. This algorithm reduces the workflows' execution costs. At the same time, the algorithm indulges the determined restriction deadline. From the results it is concluded that the proposed algorithm is promising.

Moreover, in [19] researchers talk about the networking perspectives of three popular applications. These are YouTube, Facebook, and WhatchApp. Researchers analyzed the traffic and the network infrastructure which hosts these data flows. The DBStream platform was used to analyze the large amounts of data. Solutions for traffic monitoring, analysis, and services of cellular networks have also been proposed and discussed.

The Big Data are usually transmitted from the data production center to the remote environment so that it can be provided the analysis of these large amounts of data. The multiple bandwidth reservation requests issue is discussed with the use of a High-Performance Network (HPN) in which succeeded with the best average transmission. So, in [20] have been proposed two efficient and high-speed algorithms with polynomial time complexity. The algorithms were compared with two others and from the experimental results were both verified for their advanced performance.

The pervasive network services outstretch into ubiquitous computing environment. The users to get the services they need, they have to share personal and private information. To avoid the exposure to various attacks (eg. eavesdropping) researchers proposed in [21] a security scheme to secure the communications. The authentication scheme guarantees reliability and availability by securing the remote access in Pervasive Computing Environment (PCE). The scheme provides security and convenience to the users.

8.3 Simulation Method

Based on previous works, with the aim to succeed a new type of network, which could provide more efficient data transmission a simulation tool was used. The simulation platform that used in this work is CloudSim. This simulation platform operates in the Eclipse environment, in java programming language. In CloudSim using the logic of a virtual system, and thus virtual management, we create Virtual Machines (VMs) [22] [23] [24] [25].

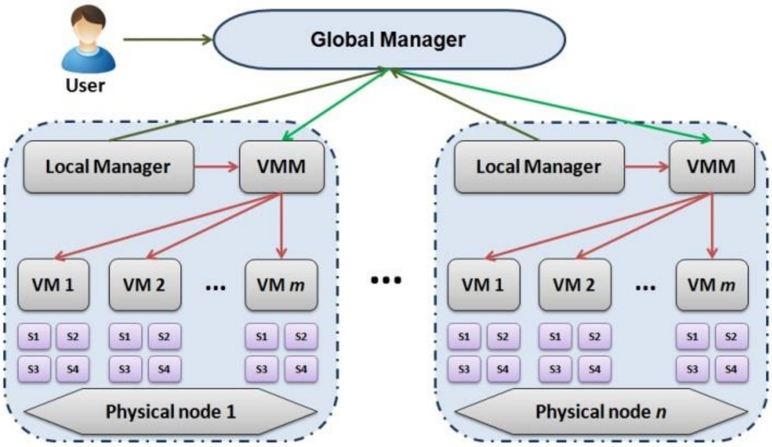


Figure 9.1: Cloud System Model

Figure 9.1 demonstrates how a user could interact through a *Global Manager* (application software) to a number of *Cloud Virtual Machines*. More specifically, each *Cloud Virtual* Machine consists of a Local Manager which interacts with a *Virtual Machine Manager* (VMM), and through the VMM established a communication path with the various individual VM devices. Each VM is connected to four sensors, from which it receives the data it then transmits to VMM. For each *Cloud Virtual Machine* there is a *Physical Node* which connects it to the network.

	Server Configuration 1	**Server Configuration 2**
Model	Dell PowerEdge T110	HP ProLiant ML110 G5
CPU Model	Intel E2160, 2C 1800MHz	Intel Xeon 3075, 2C 2660MHz
RAM	4GB	4GB
Network Bandwidth	1GB/sec	1GB/sec
Performance	1800 MIPS/core	2660 MIPS/core
Number of Servers Used	400	400

Table 9.1: Cloud Servers Configuration.

Table 9.1 lists the two types of Virtual Server Configuration for the Cloud which have been used for the simulation. In this work we used 400 Virtual Servers of the model Dell PowerEdge T110 and 400 Virtual Servers of the model HP ProLiant ML 110 G5.

Consumption in %	Dell PowerEdge T110	HP ProLiant ML110 G5
0%	86	93.7
10%	89.4	97
20%	92.6	101
30%	96	105
40%	99.5	110
50%	102	116
60%	106	121
70%	108	125
80%	112	129
90%	114	133
100%	117	135

Table 9.2: Power Consumption Information in Watt.

Table 9.2 depicts the rate of *Watt Power Consumption* from the information produced and transmitted from each type of Cloud Server, either *Dell PowerEdge T110* or *HP ProLiant G5*. As we can observe, when the rate of consumption of watts increases, both the transmission of information increases.

	VM 1	**VM 2**	**VM 3**	**VM 4**
CPU Type	High-CPU medium instance	Extra Large instance	Small instance	Micro instance
Number of Cores	1 Core	2 Cores	3 Cores	4 Cores
RAM	0.85GB	3.75GB	1.7GB	613MB
Network	1GB/sec	1GB/sec	1GB/sec	1GB/sec
Performance	2500 MIPS/core	2000 MIPS/core	1000 MIPS/core	500 MIPS/core

Table 9.3: Virtual Machine Configuration.

Table 9.3 shows the four types of VM that created and used for the simulation method. Each type had differentiated characteristics in order to be studied a wide range of results.

Length (MB)	**File size (MB)**	**Output size (MB)**
5000	5000	5000

Table 9.4: Cloudlet Parameters.

Table 9.4 demonstrates the Cloudlet Parameters which represent the volume of data used in a network in association with IoT technology. With the aim to proceed at a better simulate of high quality data, referring to digital data, we used large sizes in MB.

Subsequently, for further simulation procedure, we used *Cooja Contiki* simulator with the purpose of personalizing and extracting our network data in an environment with a defined topology.

9.4 Experimental Results

Having already tested the performance of the network we created in *CloudSim*, we perform a simulation in *Cooja Contiki*, where we tried to map the same network scenario, but also studied more aspects of this network.

We implement a network topology of a small part of the previous scenario where examining a single network segment. Namely, we examined the communication and the efficiency of data transmission in a VMM, which includes in its range five VMs.

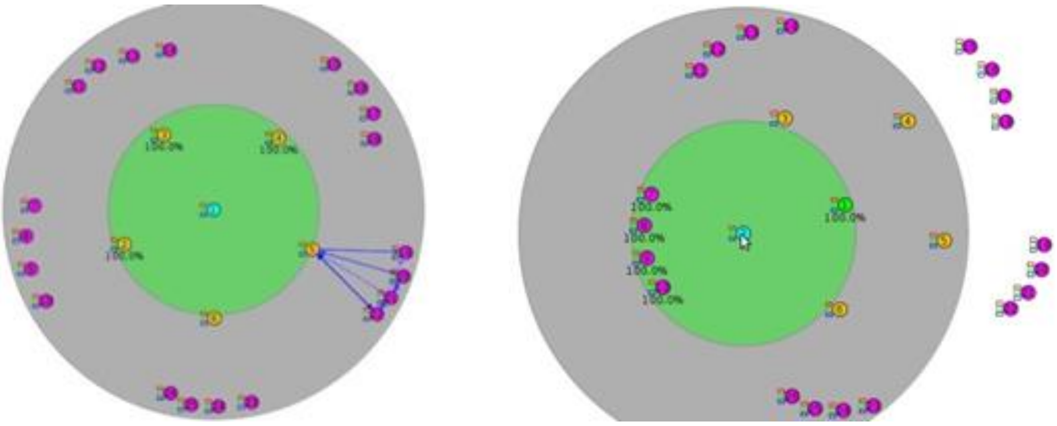


Figure 9.2: Network Topology ([2a] and [2b])

Figure 9.2a and Figure 9.2b show the topology of our proposed network. As we already mentioned, each separate part of our network consists of one VMM and five VMs. In each VM are connected four sensors. The range of the VMM contains only the five VMs, and the range of every VM contains only its four sensors. According to this, we observe that each VM contains only four sensors in its range, so as not to be inserted from the range of other VMs.

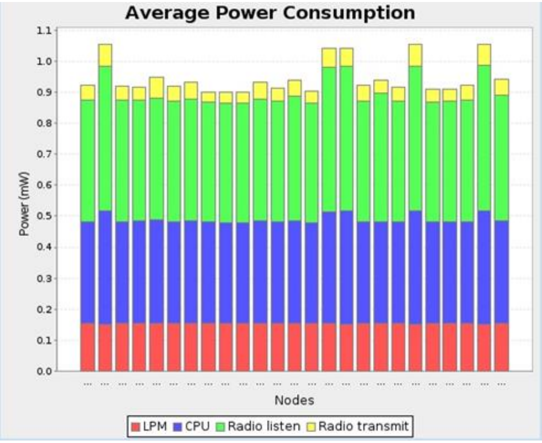


Figure 9.3a: Average Power Consumption

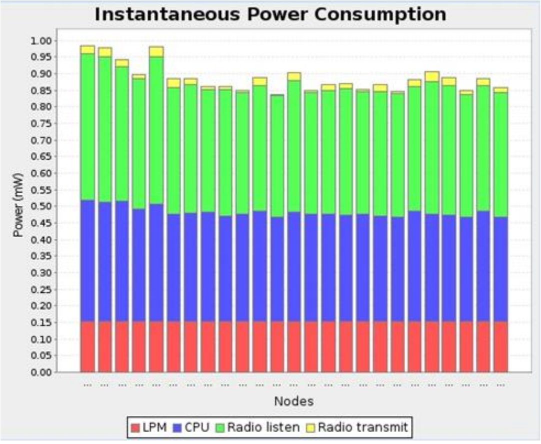


Figure 9.3b: Instantaneous Power Consumption

Figure 9.3a and Figure 9.3b demonstrates the Power Consumption of the Network. Figure 9.3a demonstrates the average Power Consumption, where we can observe that LPM’s power (red color) remains almost constant over time, as well as CPU’s power (blue color). In contrast, the Radio listen’s power (green color) and a little less the Radio transmit’s power (yellow color) where there is a greater variation in Power Consumption. Figure 9.3b demonstrates the Instantaneous Power Consumption, where, same as before, we can observe that LPM’s power (red color) remains almost constant over time, as well as CPU’s power (blue color). And also, in contrast again, the Radio listen’s power (green color) and a little less the Radio transmit’s power (yellow color) where there is a greater variation in Power Consumption. In Instantaneous Power Consumption we observe that there is a big difference as regards the variation of the Radio transmit’s power, compared with Average Power Consumption, as there are momentary fluctuations in the change in energy consumption during transmission.

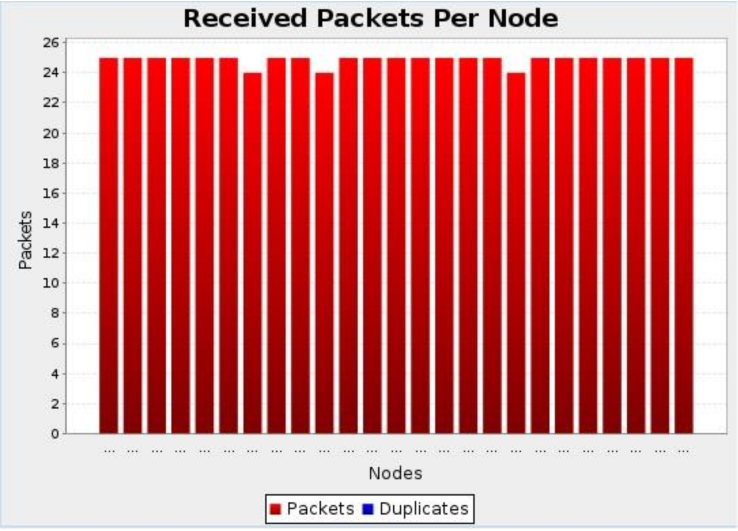


Figure 9.4: Received Packets Per Node

Figure 9.4 shows the transmitted packets which have been received per node. As we can observe, in most cases and almost every time all the nodes received the same number of packets. In addition to this, we can conclude that there are no duplicated packets received.

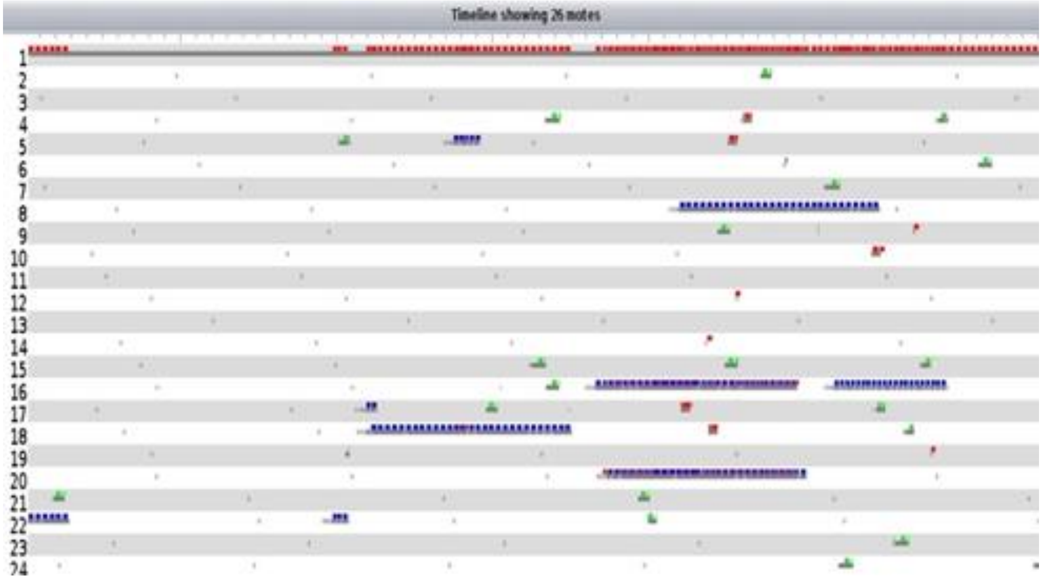


Figure 9.5: Timeline showing the packets received per mote (node) over time.

Figure 9.5 demonstrates the packet transmission procedure through all the motes we used for the simulation in *Contiki*. With the word '*mote*' is defined the node in '*Contiki language*'. Through Figure 9.5, we can see that during the simulation process transmission, there are array packages with large size (large-scale data).

9.5 Chapter Summary

Due to the blooming of IoT in CC which takes part in the last years, the there is a need of discovering new methods of technological support in many sciences by the researchers. As part of these researches, in this work, with the aim to achieve a type of network that will provide more intelligent media-data transfer, we have studied new technologies, and the use of various open source tools, such as CC analyzers and simulators. Tool like these are useful for studying the collection, the storage, the management, the processing, and the analysis of large volumes of data. Furthermore, the simulation platform used is CloudSim and operates on Eclipse environment. Thus, after measuring the network performance with CloudSim, we use the Cooja emulator of the Contiki OS in order to confirm and access more metrics and options. As a result, we implemented a network topology from a small section of the script of CloudSim with Cooja, so that we can simulate a single network segment. The results of the experiment show that there are not duplicated packets received.

Finally, as future research, we suggest a further examination of the simulation analysis of the network performance in CloudSim simulator, and other simulation platforms, with the aim to have a better and improved contribution of the technology of Internet of Things with the additional 'help' of the Cloud Computing technology for the purpose of better transmission of high quality data. This research could be a start point for better and more efficient media data transmission.

9.6 Chapter References

[14] L. Nkenyereye, J. Jang, "Design of Data Center Environmental Monitoring System Based On Lower Hardware Cost", Journal of Multimedia and Information System, Vol. 3, No. 3, pp. 63-68, October 2016.

[15] J.-H. Huh, K. Seo, "PLC-Based Smart grid Home Network System Design and Implementation using OPNET Simulation", Journal of Multimedia and Information System, Vol. 1, No. 2, pp. 111-118, December 2014.

[16] J.-H. Huh, Y. Kim, K. Seo, "Power Aware Routing Protocol in Multimedia Ad-hoc Network Considering Hop Lifetime of Node", Vol. 1, No. 2, pp. 101-110, December 2014.

[17] M. Zheng, W. Wang, "Distributed Multimedia Scheduling in the Cloud", Journal of Multimedia and Information System, Vol. 2, No. 1, pp. 143-152, March 2015.

[18] P. Fiadino, P. Casas, A. D'Alconzo, M. Schiavone, A. Baer. "Grasping Popular Applications in Cellular Netwoks with Big Data Analytics Platform", IEEE Transactions on Network and Service Management, vol. 13, issue: 3, pp. 681-695, September 2016.

[19] B. Lin, W. Guo, N. Xiong, G. Chen, A. V. Vasilakos, H. Zhang. "A Pretreatment Workflow Scheduling Approach for Big Data Applications in Multi-cloud Environments", IEEE Transactions on Network and Service Management, vol. 13, issue: 3, pp. 681-695, September 2016.

[20] L. Zuo, M. M. Zhu, "Concurrent Bandwidth Reservation Strategies for Big Data Transfers in High-Performance Networks", IEEE Transactions on Network and Service Management, vol. 12, issue: 2, June 2015.

[21] B. Djellali, P. Lorenz, K. Balarbi, A. Chouarfia. "Security Model for Pervasive Multinedia Environment", Journal of Multimedia Information System, Vol. 1, No. 1, pp. 23-43, September 2014.

Chapter 10

Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT

C. Stergiou, K. E. Psannis, B. B. Gupta, Y. Ishibashi, "Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT", Elsevier, Sustainable Computing, Informatics and Systems, vol. 19, pp. 174-184, September 2018.

10.1 Introduction

With the significant advances in communication technologies and many other sectors, also are growing up security and privacy issues. In our research, is introduced a base technology called Cloud Computing (CC) to operate with Big Data (BD). CC is a technology that refers to the processing power of data in the fog, providing more "green" computational and sustainable computing. Since it is a recently investigated technology, it has many gaps in security and privacy. So, in this paper, I proposed a new system for Cloud Computing integrated with the Internet of Things as a base scenario for Big Data. Moreover, I tried to establish an architecture relying on the security of the network to improve the security issues. A solution proposed is installing a security "wall" between the Cloud Server and the Internet, intending to eliminate the privacy and security issues. As a result, I consider that CC deals more efficiently with the privacy issue of bits transferred through time. Through the proposed system, the interaction and cooperation between things and objects communicate through the wireless networks to fulfill the objective set to them as a combined entity. Regarding the major goal of this research, which is security, a sort survey of IoT and CC presented, with a focus on the security issues of both technologies. In addition to this, I try to present the security challenges of the integration of IoT and Cloud Computing to provide an architecture relying on the security of the network to improve their security issues. Finally, I realize that through our study Cloud Computing could offer a more "green" and efficient fog environment for sustainable computing scenarios.

10.2 Related Work

For the purpose of this paper we study and analyze previous literature which has been published in the field of Big Data, CC and IoT. The following paragraphs present the papers which contributed significantly in our study.

To begin with, there are several works for the Big Data technology. In recent years several studies for BD technologies have been devised [29-34]. The authors of [29] introduce a multi-objective approach using genetic algorithms. The goal of this is to minimize two objectives, the execution time, and the budget of each node executing the task in the cloud. The contribution of [29] research is to propose an innovative adaptive model to communicate with the task scheduler of resource management. The proposed model periodically queries for resource consumption data and uses to calculate how the resources should be allocated to each task. Through this work, the authors believe that the proposed solution is timely and innovative as it provides a robust resource management where users can perform better scheduling for BD processing in a seamless manner. Furthermore, in [30] the important concepts of BD technology are highlighted and also there is a discussion about the various aspects of BD. Furthermore, the authors of [30] define what BD and discuss the various parameters of its definition. Finally, in [30] there is a look at the process that involved in the data processing and then reviewing the security aspects of BD and as a result, propose a new system for security of BD. Additionally, an offer of six provocation with the aim to spark conversations about the issue of BD technology shown in [31]. These provocations are the cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis, and mythology that dares extensive utopian and dystopian rhetoric. Finally, a multi-stakeholder approach for developing a suitable privacy regulation in the age of BD presented in [32]. This argument developed in five steps: 1) A review of the current academic debate on privacy regulation. 2) An argue that the framework for developing a suitable privacy regulation should not only focus on formal and procedural but should additionally include some important essential aspects to guard users and promote socially beneficial BD applications. 3) An examination of how the process leading to an appropriate regulation might be organized. 4) A discussion of the potential structure of a privacy organization which might conduct multi-stakeholder-dialogues as a preliminary step. 5) A discussion of their findings and suggestions.

Moreover, there are several works for the BD technology in regard with new technologies. A literature review of BD and its related technologies, such as Cloud Computing (CC) and Hadoop presented in [33]. Also, the [33] focuses on the five phases of the value chain of BD technology and as a result examines the several representative applications of BD technology. Furthermore, in [34] the important concepts of BD technology are highlighted and also there is a discussion about the various aspects of BD. Furthermore, the authors of [34] define what BD and discuss the various parameters of its definition. Finally, in [34] there is a look at process that involved in the data processing and then reviewing the security aspects of BD and as a result propose a new system for security of BD. Also, an introduction to the MIS Quarterly Special Issue on Business Intelligence Research which first offers a framework that identifies the evolution, applications and emerging research areas of BI&A presented in [35]. Moreover, a definition and description of BI&A 1.0, BI&A 2.0 and BI&A 3.0 in terms of their key characteristics and capabilities are presented in [35]. Also, there is a report of a biometric study of critical BI&A publications, researchers and research topics which based on more than a decade of related academic and industry publications are presented in [35].

As regard the Sustainability of the Cloud Computing, also, there are various works and researches that have been made in the field. We try to present those researches from oldest to newest. Initially, through the [36] the authors strive to compare and contrast Cloud Computing with Grid Computing from different angles, and in addition to give insights into the essential characteristics of both. Another research regarding the open challenges in the field presented in [37]. The [37] presents vision, and architectural elements, except for the challenges, for energy-efficient management of Cloud computing environments. The authors focus on the development of dynamic resource provisioning and allocation algorithms which consider the synergy between different data center infrastructures, and holistically work to boost data center energy efficiency and performance. More specifically, the

[37] proposes three things. At first, architectural principles for energy-efficient management of Clouds is proposed. Secondly, proposed some energy-efficient resource allocation policies and scheduling algorithms considering quality-of-service expectations, and devices power usage characteristics. And finally, the authors proposed a novel software technology for energy-efficient management of Clouds. Furthermore, the authors of [38] through their work state some major challenges in the field of Sustainable Cloud Computing, count on recent researches that have been made. One of these challenges is that it is unclear which application areas of IT can and will be outsourced to a Cloud. In a more recent work, and as a newer version of [37], the [39] defines an architectural framework and principles for energy-efficient Cloud computing. The authors, based on this proposed architecture, present their vision, open research challenges, and resource provisioning and allocation algorithms for energy-efficient management of Cloud computing environments. Additionally, the authors conduct a survey of research in energy-efficient computing and propose three things that have discussed in their past work [37]. At the end, the author of [40] discusses a thorough introduction to cloud computing which is realized with emphasis on its advantages for environmental sustainability. Also, a list of challenges in relation to the use of the technology as green technology is presented, and the reasons for using cloud computing for sustainability are explained in his work. And finally, a detailed list of the applications of Cloud Computing focusing on social, business, and environmental sustainability are listed, and a number of conclusions are provided in this work.

10.3 Security Issues in IoT & Cloud Computing Integration

There is a quick and independent evolution considering the two words of IoT and CC. Initially, the virtually unlimited capabilities and resources of CC with the aim to compensate its technological constrains, such as processing, storage and communication, could be a benefit for the Internet of Things technology. Also, the IoT technology spins out its scope to deal with real world things in a more distributed and dynamic manner and by delivering new services in a large number of real life scenarios, might be beneficial for the use of CC technology. On several occasions, CC can offer the intermediate layer between the things and the applications, hiding all the complexity and functionalities necessary to implement the latter [41].

Through the integration of IoT and CC could be observed that CC can “complete” some gaps of IoT, such the “*limited storage*” and “*applications over internet*”. Also, IoT can “complete” some gaps of CC, such the main issue of “*limited scope*”. Based on motivations such those referred beforehand, and the important issue of security in both technologies we can assume some motivations for the integration. The security issue of this integration has a serious problem. When critical IoT applications move towards the CC technology, concerns arise due to the lack of trust in the service provider or the knowledge about service level agreements (SLAs) and knowledge about the physical location of data. Consequently, new challenges require specific attention as mentioned in surveys [42] [43]. Multi-tenancy could additionally conciliate security and lead to sensitive information leakage. Furthermore, public key cryptography cannot be applied at all layers due to the computing power constraints imposed by the things. These are examples of topics that are currently under investigation with the aim to tackle the big challenge of security and privacy in CC and IoT integration [41].

Subsequently, some challenges about the security issue in the integration of two technologies are listed below [42]:

a) *Heterogeneity*. A big challenge in CC and IoT integration is related to the wide heterogeneity of devices, operating systems, platforms, and services available and possibly used for new or improved applications [44].

b) *Performance*. Often CC and IoT integration's applications introduce specific performance and QoS requirements at several levels and in some particular scenarios meeting requirements may not be easily achievable [45].

c) *Reliability*. When Cloud Computing and IoT integration is adopted for mission-critical applications, reliability concerns typically arise. When applications are developed in resource constrained environments several challenges related to device failure or not always reachable devices exists [46].

d) *Big Data*. With an estimated number of 50 billion devices that will be networked by 2020, specific attention must be paid to transportation, storage, access, and processing of the huge amount of data they will produce [47].

e) *Monitoring*. As largely documented in the literature, monitoring is an essential activity in CC environments for capacity planning, for managing resources, SLAs, performance and security, and for troubleshooting [48].

IoT & Cloud Computing security challenges	**Heterogeneity**	**Performance**	**Reliability**	**Big Data**	**Monitoring**
Internet of Things		**X**	**X**	**X**	**X**
Cloud Computing	**X**	**X**		**X**	

Table 10.2: Affects of IoT & Cloud Computing security challenges.

Table 10.2 lists the two abovementioned technologies that have studied in this research and the challenges of their integration which arising from our study. These challenges are related to the security issue in the integration of those two technologies. As we can observe from Table 10.2, the both technologies have two common main challenges of their integration which are *Performance* and *Big Data*. Additionally, we can realize that IoT technology related to more challenges (4) than the CC technology (3).

10.4 Proposed System

The study of previous works cites us relevant architecture and topology proposals for a Smart Building network, which on several occasions supported and operated in Internet of Things and Fog environments. In this section we will make a comparative analysis study of the some previous works which we have distinguished. Initially, we analyze what each of them deals with.

Regarding the Literature Review analysis we realize that not enough works deal with security and privacy issues in Cloud Computing for technologies such as Big Data and Internet of Things. Thus, we try to propose a new system for Cloud Computing integrated with Internet of Things as a base scenario for Big Data. In order to improve the security issues we would try to establish an architecture relaying on the security of the network.

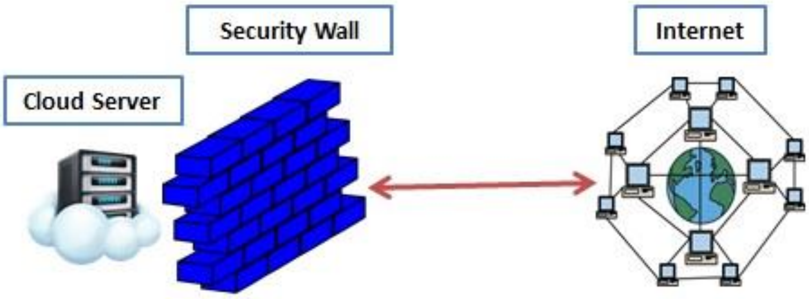


Figure 10.1: Network Scenario based in "strong" Security Wall.

As shown in Figure 10.1, a security "wall" installed between the Cloud Server and the Internet (the various users), with the aim to eliminate the privacy and security issues. This type of network uses all the benefits of the existing topologies (e.g. star, ring etc.) in order to have better communication and to transfer more safely large-scale data (Big Data) through the network.

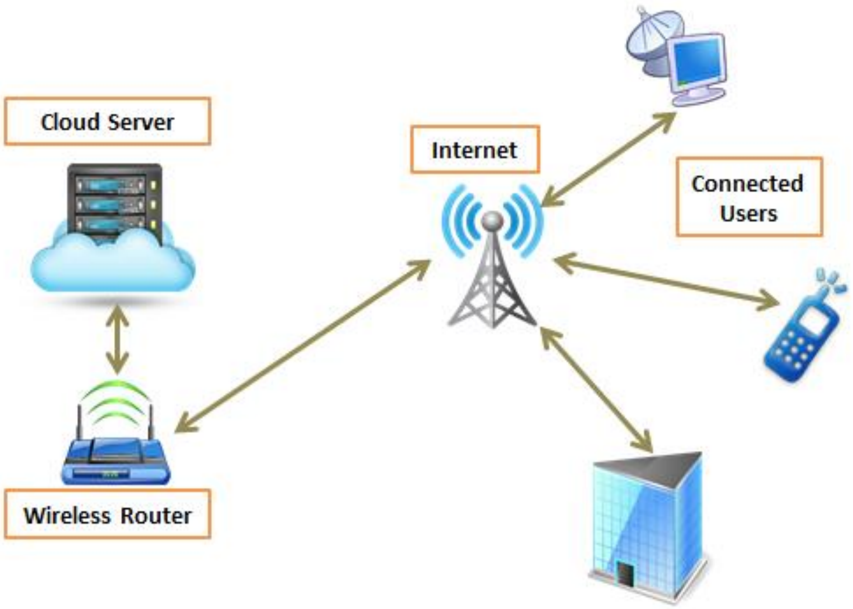


Figure 10.2: Wide-Range Network of Cloud Computing.

By applying the proposed model we can extend the advances of IoT and Cloud Computing, by developing a highly innovative and scalable service platform to enable secure and privacy services. Through our research which carried out we can propose the following part of algorithm which extends the security advances of both Cloud and IoT technologies. As a proposal of this work could be this part of algorithm which uses the original key consists of 128 bits/16 bytes which are represented as a 8x8 matrix, represented bellow.

Algorithm 1

```
Cipher(byte[] input, byte[] output)
{
    byte[8,8] State;
    copy input[] into State[] AddRoundKey
     for (i=8; i<88; i++)
     {
        T = W[i-1];
        if (i mod 8 = = 0)
            T = Substitute (Rotate (T)) XOR RConstant [i/8];
        W[i] = W[i-4] XOR T;
        SubBytes ShiftRows MixColumns
     AddRoundKey
     }
    SubBytes ShiftRows AddRoundKey
    copy State[] to output[]
}
```

The new system architecture provides safer paths between the users of the network. As we can see in Figure 10.2, the whole connection relaying on wireless communication. The Server connects to the internet through a simple wireless router, where there is installed the “security wall”. Through the internet any type of user could have access and manage the transferred data of the network if it meets the requirements.

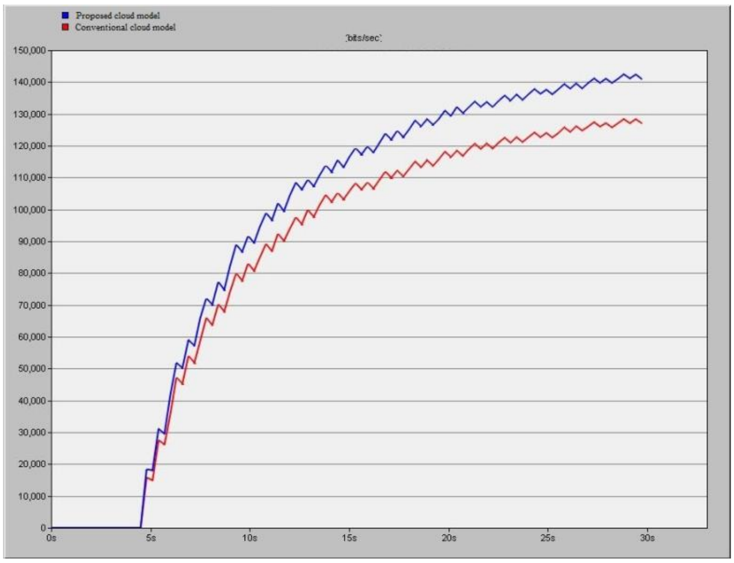
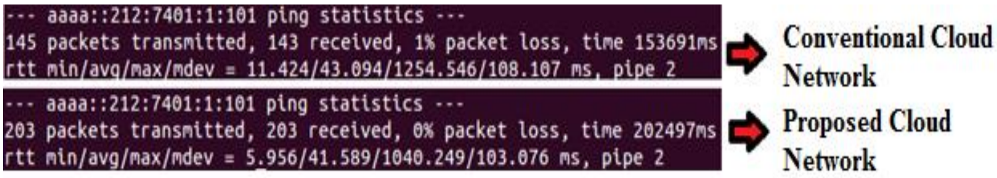


Figure 10.3: Efficiency comparison proposed model vs. conventional Cloud model.

Figure 10.3 demonstrates the "*little*" difference between the two models that offers remote access and management of data. As we can see, our proposed model is a little bit better dealing with the privacy issue of bits transferred through time. With the use of Wireshark we test the packets sent and received in our proposed Cloud network and in a conventional Cloud network with the same configuration. The packet loss in the conventional Cloud network is a little bit more in contrast with the our proposed Cloud network. This is demonstrated in Figure 10.4 below:

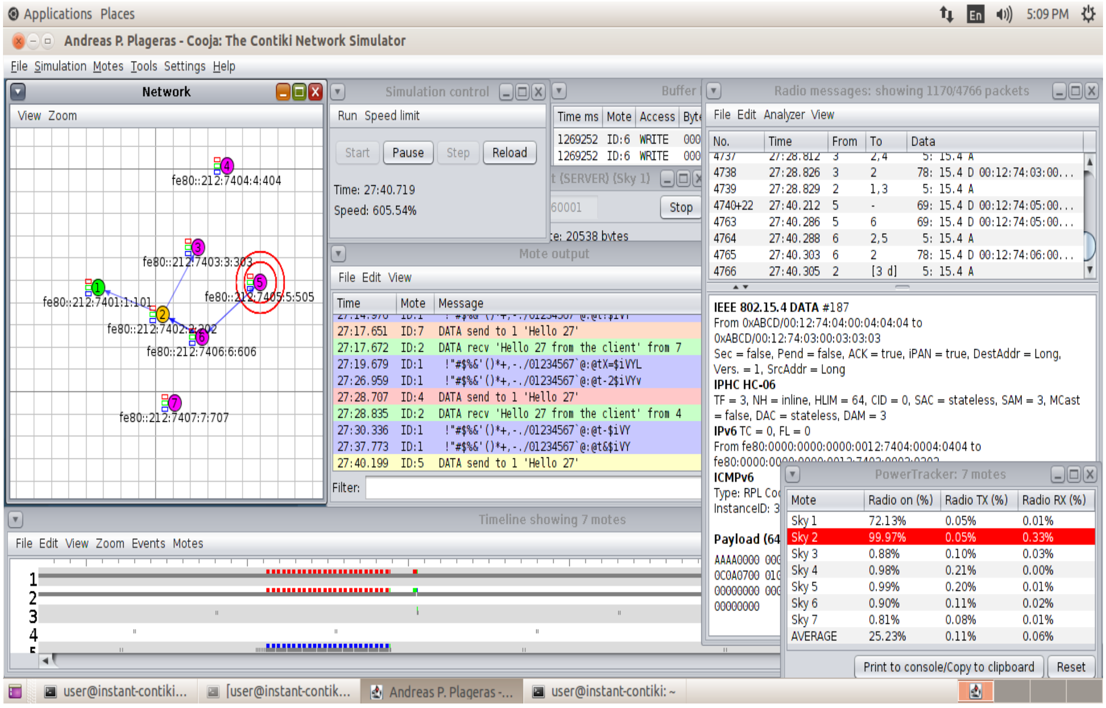


```
--- aaaa::212:7401:1:101 ping statistics ---
145 packets transmitted, 143 received, 1% packet loss, time 153691ms
rtt min/avg/max/mdev = 11.424/43.094/1254.546/108.107 ms, pipe 2
```

```
--- aaaa::212:7401:1:101 ping statistics ---
203 packets transmitted, 203 received, 0% packet loss, time 202497ms
rtt min/avg/max/mdev = 5.956/41.589/1040.249/103.076 ms, pipe 2
```

Figure 10.4: The packet loss for both conventional cloud and proposed cloud computing network respectively.

where the first output from the ping we made is for the conventional Cloud scenario and the second output from the ping we made is for our proposed Cloud scenario. We used the Contiki operating system and the Cooja emulator for the simulation of the edge computing network. The network and various information about it, are presented in Figure 10.5 below.



Time: 27:40.719
Speed: 605.54%

Time ms	Mote	Access	Byte
1269252	ID:6	WRITE	000
1269252	ID:6	WRITE	000

te: 20538 bytes

```
Time       Mote  Message
27:17.651  ID:7  DATA send to 1 'Hello 27'
27:17.672  ID:2  DATA recv 'Hello 27 from the client' from 7
27:19.679  ID:1   !"#$%&'()*+,-./01234567`@:@tX=$iVYL
27:26.959  ID:1   !"#$%&'()*+,-./01234567`@:@t-2$iVYv
27:28.707  ID:4  DATA send to 1 'Hello 27'
27:28.835  ID:2  DATA recv 'Hello 27 from the client' from 4
27:30.336  ID:1   !"#$%&'()*+,-./01234567`@:@t-$iVY
27:37.773  ID:1   !"#$%&'()*+,-./01234567`@:@t&$iVY
27:40.199  ID:5  DATA send to 1 'Hello 27'
```

```
No.      Time       From  To     Data
4738     27:28.826  3     2      78: 15.4 D 00:12:74:03:00...
4739     27:28.829  2     1,3     5: 15.4 A
4740+22  27:40.212  5     -      69: 15.4 D 00:12:74:05:00...
4763     27:40.286  5     6      69: 15.4 D 00:12:74:05:00...
4764     27:40.288  6     2,5     5: 15.4 A
4765     27:40.303  6     2      78: 15.4 D 00:12:74:06:00...
4766     27:40.305  2     [3 d]   5: 15.4 A
```

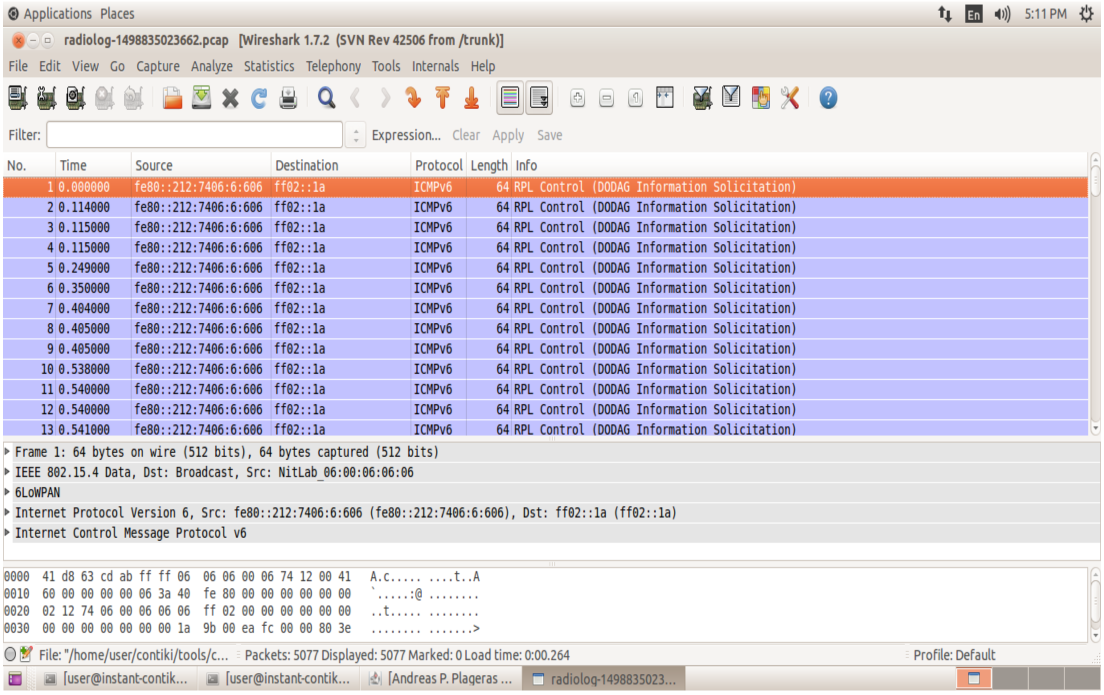
IEEE 802.15.4 DATA #187
From 0xABCD/00:12:74:04:00:04:04:04 to
0xABCD/00:12:74:03:00:03:03:03
Sec = false, Pend = false, ACK = true, iPAN = true, DestAddr = Long,
Vers. = 1, SrcAddr = Long
IPHC HC-06
TF = 3, NH = inline, HLIM = 64, CID = 0, SAC = stateless, SAM = 3, MCast
= false, DAC = stateless, DAM = 3
IPv6 TC = 0, FL = 0
From fe80:0000:0000:0000:0012:7404:0004:0404 to
fe80:0000:0000:0000:0012:7403:0003:0303
ICMPv6
Type: RPL Co
InstanceID: 3

Payload (64
AAAA0000 000
0C0A0700 010
00000000 000
00000000

Mote	Radio on (%)	Radio TX (%)	Radio RX (%)
Sky 1	72.13%	0.05%	0.01%
Sky 2	99.97%	0.05%	0.33%
Sky 3	0.88%	0.10%	0.03%
Sky 4	0.98%	0.21%	0.00%
Sky 5	0.99%	0.20%	0.01%
Sky 6	0.90%	0.11%	0.02%
Sky 7	0.81%	0.08%	0.01%
AVERAGE	25.23%	0.11%	0.06%

Figure 10.5: The simulation of the proposed cloud computing network using the Contiki OS.

Furthermore, in Figure 10.6, are shown the packets collected which are stored by a "radio messages" tool in .pcap files for further analysis. Wireshark is a network sniffer and analyzer in which, we can observe in the following Figure 10.6, various information about the communications.


Applications Places
5:11 PM
radiolog-1498835023662.pcap [Wireshark 1.7.2 (SVN Rev 42506 from /trunk)]
File Edit View Go Capture Analyze Statistics Telephony Tools Internals Help
Filter:
Expression... Clear Apply Save
Profile: Default
[user@instant-contik...
[user@instant-contik...
[Andreas P. Plageras ...
radiolog-1498835023...

```
No. Time      Source               Destination  Protocol Length Info
  1 0.000000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  2 0.114000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  3 0.115000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  4 0.115000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  5 0.249000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  6 0.350000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  7 0.404000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  8 0.405000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
  9 0.405000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
 10 0.538000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
 11 0.540000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
 12 0.540000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)
 13 0.541000  fe80::212:7406:6:606 ff02::1a     ICMPv6   64 RPL Control (DODAG Information Solicitation)

▸ Frame 1: 64 bytes on wire (512 bits), 64 bytes captured (512 bits)
▸ IEEE 802.15.4 Data, Dst: Broadcast, Src: NitLab_06:00:06:06:06
▸ 6LoWPAN
▸ Internet Protocol Version 6, Src: fe80::212:7406:6:606 (fe80::212:7406:6:606), Dst: ff02::1a (ff02::1a)
▸ Internet Control Message Protocol v6

0000  41 d8 63 cd ab ff ff 06  06 06 00 06 74 12 00 41   A.c..... ....t..A
0010  60 00 00 00 00 06 3a 40  fe 80 00 00 00 00 00 00   `.....:@ ........
0020  02 12 74 06 00 06 06 06  ff 02 00 00 00 00 00 00   ..t..... ........
0030  00 00 00 00 00 00 00 1a  9b 00 ea fc 00 00 80 3e   ........ .......>

File: "/home/user/contiki/tools/c... Packets: 5077 Displayed: 5077 Marked: 0 Load time: 0:00.264
```

Figure 10.6: The packets collected and analyzed with Wireshark.

There are more tools, like Wireshark, available to be used with the Contiki OS, such as foren6 and other packet sniffers and packet analyzers. Also, with the Contiki OS we can handle the security and the privacy of these data collected, but novel ideas and implementations need to be further investigated in order to simulate such solutions in real-time and efficiently.

10.5 Experimental Results

Through experimental scenarios which we have made we have strengthened our suggestion that our proposed model is more efficient than the conventional Cloud model. We perform a number of simulations and measurements through which we can realize that we have done a good effort.

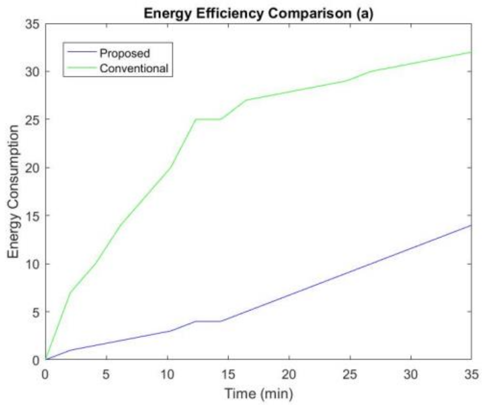


Figure 10.7: Energy Efficiency Comparison (a).

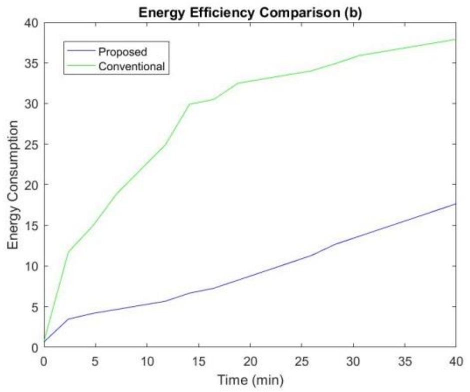


Figure 10.8: Energy Efficiency Comparison (b).

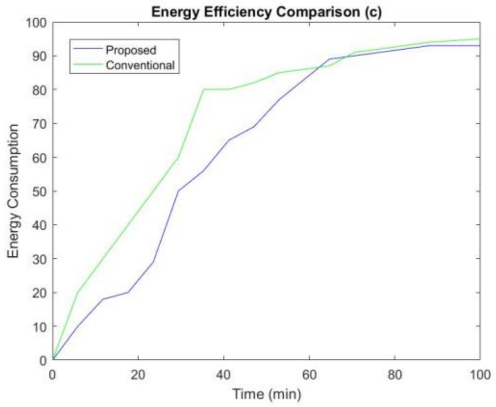


Figure 10.9: Energy Efficiency Comparison (c).

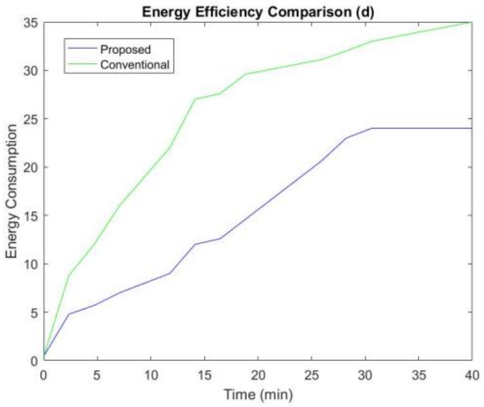


Figure 10.10: Energy Efficiency Comparison (d).

Figure 10.7, Figure 10.8 Figure 10.9 and Figure 10.10 describe the better Energy Efficiency offered in a framework implementing by our proposed model. As we can observe our proposed model needs less energy power than the conventional model through the time. This can offer a better option of energy consumption and provide a more environmental friendly framework.

The four Figures above demonstrate the difference between the two models that offers remote access and management of data. As we can see, through the experimental results, our proposed model is a bit better dealing with the privacy issue of bits transferred through time.

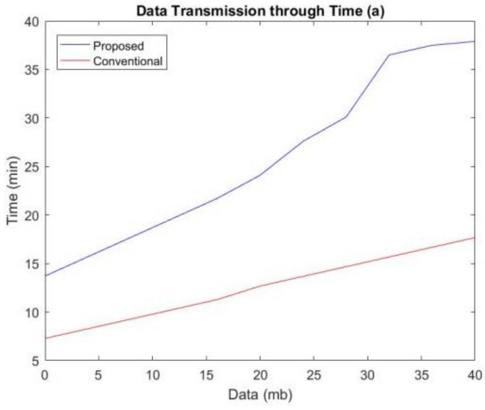


Figure 10.11: Data Transmission through Time (a).

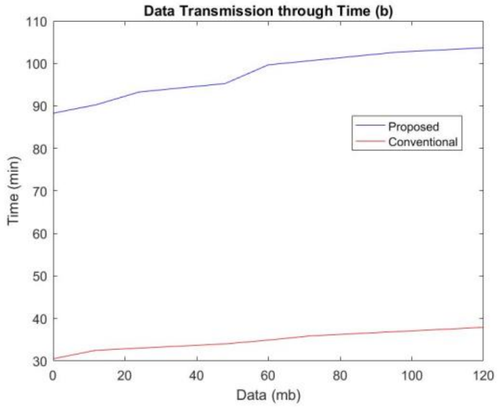


Figure 10.12: Data Transmission through Time (b).

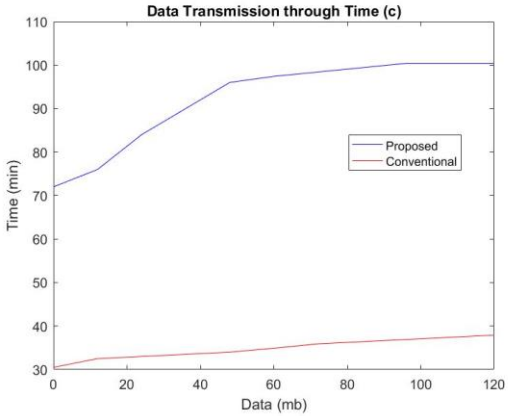


Figure 10.13: Data Transmission through Time (c).

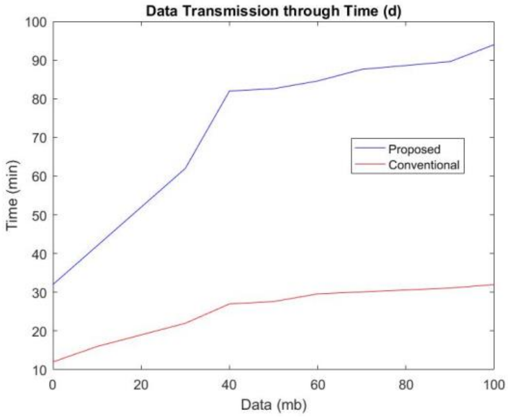


Figure 10.14: Data Transmission through Time (d).

Moreover, through Figure 10.11, Figure 10.12, Figure 10.13 and Figure 10.14 we can observe that through our proposed model we can achieve higher data transmission rate through the time than the conventional model. Due to our efficient settled network our model can transmitted higher amounts of data through time, when the conventional model can transmit fewer amounts of data.

10.6 Advantages and Benefits of the Proposed Model

Cloud Computing could offer many benefits to people in general, businesses and Small and Medium Enterprises in particular through our proposed model, but additionally and in general use. The five main reasons for adopting a Sustainable Computational Cloud technology with the aim to give it extra boost and competitiveness are listed below:

- ✓ Offers software and application solutions without greatly increasing costs as applications run on the Cloud and businesses do not need expensive computing systems.
- ✓ Has in a similar size data storage in the Cloud proportional to the package selected by the customer.
- ✓ Gives access to Cloud data from anywhere and any device at any given time, giving portability and flexibility to the business.
- ✓ It is backed by state-of-the-art security protocols that ensure enterprise data protection.
- ✓ Provides optimal business performance due to flexibility, mobility and productivity.

Regarding the efficiency of Cloud Computing in general, and more specific our proposed model, there are in addition economic and efficiency benefits:

- ✓ Reduces labor costs by 50% in the configuration, operation, monitoring and management of business operations.
- ✓ Improves the quality and elimination of software defects by up to 30%.
- ✓ Reduces support costs for end users up to 40%.

All these could consist that Cloud is a “green” computational scenario.

10.7 Chapter Summary

Through our research we found in the conclusion that security and privacy issues grew up by the significant advances in the sector of communications and additionally in other sectors. Relay on this, this work aims to introduce Cloud Computing as a base technology in order to operate and integrate with recent technologies such as Big Data and Internet of Things. The technology of Cloud Computing refers to the processing power of the data at the “edge” of a network. Additionally, we could say that Cloud Computing operates in “fog” environment. Regarding this and relaying on the privacy issues in its operation, we proposed a new system for CC which integrated with IoT and operates as a base scenario for BD.

In our scenario, we tried to make an establishment of an architecture relaying on the security of the network with the aim to improve the security and privacy issues. Thus, through the simulations which have been made, the solution that we proposed installs a security “wall” between the Cloud Server and the different users in the Internet. This proposal aims to eliminate the privacy and security issues which need to be faced. Concluding, we considered that the CC provides efficiency in privacy issues of the network, where bits transferred through time.

The main goal of the interaction and cooperation between things and objects communicate through the wireless networks, in order to fulfil the objective set to them as a combined entity. In this research, a sort survey of IoT and CC was presented, with a focus on the security issues of both technologies. Furthermore, the security challenges of the integration of IoT and Cloud Computing were surveyed through the proposed architecture. At the end, we survey the security challenges of the integration of IoT and Cloud Computing with the aim to provide an architecture relaying on the security of the network in order to improve the security issues. Concluding, we could state that Cloud Computing could offer a more "green" and efficient fog environment for sustainable computing scenarios.

Regarding the future, we plan to make more simulations in order to have a better accuracy in our experimental results. More data transfer scenarios have to be made through the simulators providing results counting not only in data transmission but also in network efficiency and support. Also, considering that CC is a novel technology in the sector of communications, more study need to be made about its operation and how CC interact and integrates in better way with other technologies such as IoT and BD. This can be the field of future research.

10.8 Chapter References

[29] I. A., T. Hashem, N. B. Anuar, A. Gani, "Schedule optimization for big data processing on cloud", in Proceedings of 2nd International Conference on Big Data Analysis and Data Mining, San Antonio, USA, 30 November - 1 December 2015.

[30] R. Toshniwal, K. G. Dastidar, A. Nath, "Big Data Security Issues and Challenges", International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 2, issue: 2, pp. 15-20, February 2015.

[31] K. Crawford, D. Boyd, "Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon", Information, Communication & Society Journal, vol. 15, issue: 5, pp. 662-679, May 2012.

[32] M. G. Will, "Privacy and Big Data: The Need for a Multi-Stakeholder Approach for Developing an Appropriate Privacy Regulation in the Age of Big Data", Ingo Pies, Halle, July 2015 [Available at SSRN: https://ssrn.com/abstract=2634970].

[33] I. Chebbi, W. Boulila, I. R. Farah, "Big Data: Concepts, Challenges and Applications", Springer International Publishing, Computational Collective Intelligence, vol. 2, p. 638–647, October 2015.

[34] R. Toshniwal, K. G. Dastidar, A. Nath, "Big Data Security Issues and Challenges", International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 2, no. 2, pp. 15-20, February 2015.

[35] H. Chen, R. H. L. Chiang, V. C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact", MIS Quarterly, vol. 36, no. 4, pp. 1165-1188, December 2012.

[36] I. Foster, Y. Zhao, I. Raicu, S. Lu, “Cloud Computing and Grid Computing 360-Degree Compared”, in Proceedings of IEEE GCE ’08, Grid Computing Environments Workshop, Austin TX, USA, 12-16 November 2008.

[37] R. Buyya, A. Beloglazov, J. Abawajy, “Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges”, in Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2010), Las Vegas, USA, July 12-15, 2010.

[38] G. Muller, N. Sonehara, I. Echizen, S. Wohlgemuth, “Sustainable Cloud Computing”, Business & Information Systems Engineering, vol. 3, issue 3, pp. 129-131, June 2011.

[39] A. Beloglazov, J. Abawajy, R. Buyya, “Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing”, Future Generation Computer Systems, vol. 28, issue 5, pp. 755-768, May 2012.

[40] K. Domdouzis, “Sustainable Cloud Computing”, Book section in DASTBAZ, Mohammad, PATTINSON, Colin and AKHGAR, Babak, Green information technology: a sustainable approach, Waltham, MA, Morgan Kaufmann, pp. 95-110, December 2015.

[41] A. Botta, W. de Donato, V. Persico, A. Pescape, “Integration of Cloud Computing and Internet of Things: a Survey”, Journal of Future Generation Computer Systems, vol. 56, pp. 1-54, March 2016.

[42] T. Bhattasali, R. Chaki, N. Chaki, “Secure and trusted cloud of things”. In Proceedings of Annual IEEE India Conference (INDICON 2013), Mumbai, India, 13-15 December 2013.

[43] Y. Simmhan, A. G. Kumbhare, B. Cao, V. Prasanna, “ An analysis of security and privacy issues in smart grid software architectures on clouds”, In Proceedings of IEEE International Conference on Cloud Computing (CLOUD 2011), pp. 582–589, Washington, DC, USA, 4-9 July 2011.

[44] N. Grozey, R. Buyya, “Inter-cloud architectures and application brokering: taxonomy and survey”, Software: Practice and Experience, vol. 44, issue: 3, pp. 369–390, December 2012.

[45] B. P. Rao, P. Saluia, N. Sharma, A. Mittal, S. V. Sharma, “Cloud computing for Internet of Things & sensing based applications”, In Proceedings of IEEE 6th International Conference on Sensing technology (ICST 2012), pp. 374–380, Kolkata, India, 18-21 December 2012.

[46] W. He, G. Yan, L. D. Xu, “Developing vehicular data cloud services in the IoT environment”, IEEE Transactions on Industrial Informatics, vol. 10, issue: 2, pp. 1587–1595, May 2014.

[47] C. Dobre, F. Xhafa, “Intelligent services for big data science”. Elsevier, Future Generation Computer Systems, vol. 37, pp. 267–281, July 2014.

[48] G. Aceto, A. Botta, W. de Donato, A. Pescape, “Cloud monitoring: A survey”, Elsevier, Computer Networks, vol. 57, issue: 9, pp. 2093–2115, June 2013.

Chapter 11

IoT-based Big Data secure management in the Fog over a 6G Wireless Network

C. L. Stergiou, K. E. Psannis, B. B. Gupta, "IoT-based Big Data secure management in the Fog over a 6G Wireless Network", IEEE Internet of Things Journal, In Press, 2020. [DOI: 10.1109/JIOT.2020.3033131]

11.1 Introduction

This work proposes an innovative infrastructure of secure scenario which operates in a wireless-mobile 6G network for managing Big Data on Smart Buildings. Count on the rapid growth of the telecommunication field new challenges arise. Furthermore, a new type of wireless network infrastructure, the sixth generation (6G), provides all the benefits of its past versions and also improves some issues which its predecessors had. In addition, relative technologies to the telecommunications field, such as IoT, Cloud Computing, and Edge Computing, can operate through a 6G wireless network. Take into account all these, we propose a scenario that tries to combine the functions of the Internet of Things with Cloud Computing, Edge Computing, and Big Data to achieve a Smart and Secure environment. The major purpose of this work is to create a novel and secure Cache Decision System in a wireless network that operates over a Smart Building, which will offer the users a safer and efficient environment for browsing the internet, sharing and managing large-scale data in the fog. This CDS consisted of two types of servers, one Cloud Server and one Edge Server. To come up with the proposal, I study related cache scenario systems which are listed, presented, and compared in this work.

11.2 Related Work

For the purpose of the research existing literature, on the fields of 6G Networks BD, IoT, CC, EC, and caching, has been also studied and analyzed. We realized that the usage of remote storage space become more popular and widely used over the years. So, new ways in caching systems appear in order to improve the current status so far. Subsequently, there is a need for further research in this area as the growing numbers studied indicate. Figure 11.1 demonstrates the rapid growth of the papers related to the issues of cache scenarios systems, and also how other technologies, such as those mentioned in this work, can help improve it. As we can observe, the last seven years the interest of the researchers has growth considerably compared to the previous decade.

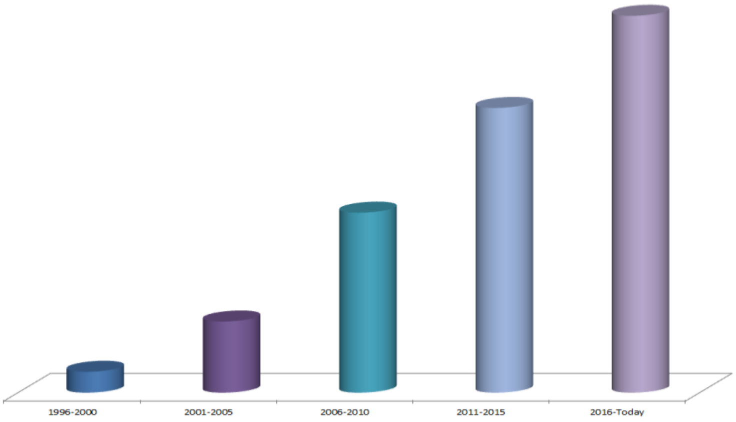


Figure 11.1: Change in the number of related works over time.

One of the biggest challenges in big data is data mining. Raichura & Padhariya in their work [20] try to deliver data in an efficient way. Specifically, to access the data efficiently and to transfer the data in a cost-effective way, researchers pro-posed a cache-based architecture, which was named "*BigCache*". Through that system, the caching of the static vs the dynamic data, the storage and the access speed have been taken into consideration and have been improved.

Also, improvements have been done in bandwidth utilization, in the availability of the data, in the updating of the data by fresh ones', and in costs which are reduced. In *BigCache* are included three caching schemes, a Cache Decision System (CDS), and a cache replacement process. As results of the performance evaluation of the system are the reduced bandwidth and the reduced server-load.

Jin et al. [18] in their work present several mobile wireless caching frameworks such as those based on Information-Centric Networking (ICN), on Mobile Cellular Networks (MCNs), on Wireless Ad Hoc Networks (WAHNs), and on Hybrid Networks (HNs). The ICN is based on name contents, which means that the network routes the users' requests by the content name. After the searching for the content in various caches, the requested content is forwarded to the users. Some technologies used in ICN are the Named Data Networking (NDN) or name based routing, the information naming, and others. In the MCNs the caching is done in both the core network and the access networks. Generally, the users through their mobile devices are connecting to the specific Access Point (AP) to gain multimedia services. In the WAHNs there are multi-hop communications by the mobile nodes, and by that fact cache can be deployed, so that the exchange of data between these nodes can be managed. Finally, in the HNs a combination of a WAHN with an infrastructural network is done. The HNs are also supporting ad hoc and cellular communications based on mobile devices.

Another region based approach discussed by Sun et al. [22] about the BD in mobile social networks and the QoE that pro-vides to the users. The big amounts of data as it is known suffer from their large volume and variety, and from their large value and velocity. Because of this, a framework for efficient big data transmission through content-centric mobile social networks to the mobile devices is presented.

Meoni et al. [23] through their work try to investigate how leverage machine learning on this huge amount of data with the aim to discover patterns and correlations useful to growth the overall efficiency of the distributed infrastructure in terms of CPU utilization and task completion time. More specifically, this work proposes a scalable pipeline of components built on top of the Spark engine for large-scale data processing. The goal of this proposal is collecting from different sites the dataset access logs then organizing them into weekly snapshots, and training, on these snapshots, predictive models which are able to forecast which datasets will become popular over time.

In their work, Djemaiel et al. [24], in the context of Wireless Sensor and Actuator Networks (WSANs), proposes a cloud-based WSAN approximation that enables the storage and the management of health data in an efficient way by representing the collected data and their reliance using Temporal Conceptual Graphs (TCGs). The authors with the aim to demonstrate the efficiency of the proposed approach illustrate for different defined scenarios of diseases and their associated health data demonstrated through generated TCGs.

At the end, Rafique et al. [25] with the aim to address challenges related to Cloud scenarios, present PERSIST, which is middleware architecture. This architecture initially externalizes the complexity of federated cloud storage architecture and the complex storage logic from the SaaS application to storage policies, allow tenants to enforce different storage grained level. Also, the PERSIST supports the dynamic re-configurability of the underlying federated Cloud storage architecture.

11.3 Evaluation Approach System-Framework

The purpose of caching is to perform better with the websites and to save bandwidth. By tacking in advance all the previous works and all the aspects that come out from the limitations in wireless networks, we would try to propose a new caching system for a wireless network. This proposed caching system is established on a SB over a University, in order to let the mobile users in the building access the internet faster and safer. The proposed system consists of a main local server which is connected to two storage cache servers, one Cloud cache server and one Edge cache server. But, how a cache server works? A cache server can be divided into two parts: the webserver cache and the application cache.

Figure 11.2 demonstrates the operation of our proposed system. More specific, the users will be connected to the internet through a mobile network settled in the campus and based on the novel 6G technology and its' benefits. The 6G network could offer to the users faster and responsive network environment. In addition to this, they can send requests to the cache servers depending the occasion. More particular, the scenario is as follows, if all users request the same content-item, then the first request is only computed and the rest users will access the requested content-item from cache since it is stored there, after the first computation. As also shown in Figure 11.2, there is a need of a proxy cache server between the local server and the webservers. Thus, thereafter the requests will be sent to the proxy first.

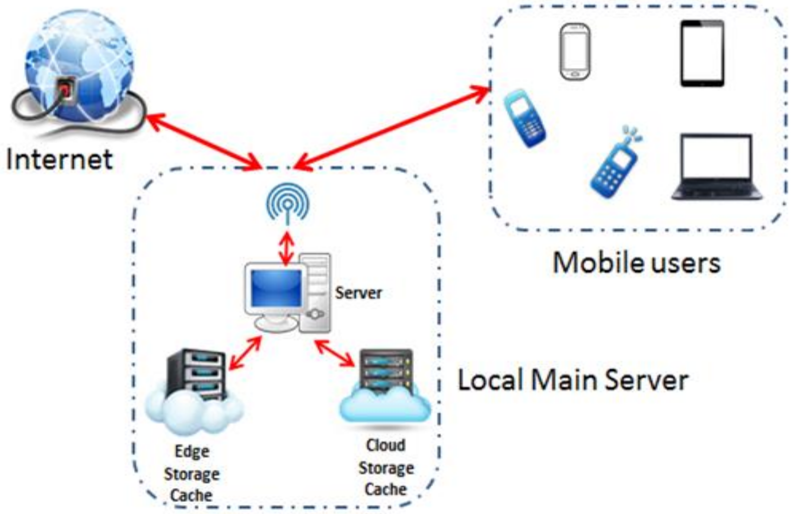


Figure 11.2: Proposed System Architecture.

More specifically, regarding the operation of the Cache Decision System, when a user tries to connect to a web page, an algorithm will run and check if the user chooses a secure web site or not. Particularly, a Cache Decision Algorithm (CDA) manages the preference and then the Mail Server retrieves data from the proper Storage Server. For instance, an example scenario is as follows, if the user wants to access a web site that was hosted as "*https://...*" the Cloud Storage Cache will be used, otherwise, if the user wants to access a web site that was hosted as "*http://...*" the Edge Storage Cache will be used. In the first situation, caching of https can be controlled with the use of response headers.

Moreover, the users which request a content-item are being categorized in those who need a secure communication and request websites under the "*https://*", and those who need speed in their communications and request daily-common websites under the "*http://*". The first category has the priority to keep its users safe. Then, the requests of "*http://*" websites can be served faster with the use of the Edge Cache Server.

Consequently, we choose Cloud Storage Server in order to store secure data because Cloud technology offers better security scenarios instead of Edge technology. On the other hand, Edge technology could offer users easier and quicker access in not securely encrypted web sites.

Based in previous works we realized that CC and EC differ in data management. CC offers better security and privacy than other technologies, such as IoT, due to its data encryption system [9] [12] [18] [26] [27]. CC used as a based technology for a major number of IoT-based and BD-based applications because of its function. On the other hand regarding previous works EC used in high-speed networks due to its feature of high-speed transmission [8] [12] [15] [28] [29] [30].

In addition to this, a CDS that is able to decide and choose the better cache system which also could manage data consisted of large-scale data and produced be various IoT devices could be count on our proposal. IoT's features such as *Service over Internet*, *Application over Internet*, *Energy Efficiency* and *Computationally Capable* could have a positive impact from the proposed system due to the use of both CSe and Ese. Furthermore, BD that transferred and used through the network of the proposed system could be stored depending the occasion either on CSe or Ese, in order to be managed from the users that have access on them. The ability of storage and services offers new possibilities of using data analysis application.

11.3.1. Scheme Implementation Scenario

After studying various researches about caching in wireless mobile networks and after we proposed a CDS for a Smart Building, it is time to consider which caching algorithm fits the proposed system. There are lots of algorithms for caching but only few are the most used and fit to such systems like the proposed one. Some of them are explained below:

- ✓ First In First Out (FIFO): the cache behaves as a FIFO queue.
- ✓ Last In First Out (LIFO): the cache behaves in the opposite way of FIFO.
- ✓ Least Frequently Used (LFU): one of the most commonly used caching algorithm. This algorithm is used to remember the most frequently used content-items. The rest, which are rarely used, are rejected first.
- ✓ Least Recently Used (LRU): one of the most commonly used caching algorithm. This algorithm is used to reject the LRU content-items first. It is also used by Dropbox and Android.
- ✓ Least Frequent Recently Used (LFRU): A hybrid algorithm of the previous two: LFU and LRU.
- ✓ Most Recently Used (MRU): This algorithm works in the opposite way than the LRU algorithm. This means that the rarely used content-items are kept in cache and the recently used are rejected.
- ✓ Adaptive Replacement Cache (ARC): This algorithm is a combination of LFU and LRU and it improves the result of that combination. It also makes a good management of the cache storage spaces.

The algorithm that fits to the proposed system is the LFRU. This hybrid algorithm is divided into two partitions. The first one which is called "*privileged partition*" (where it is used the LRU algorithm) is used for the popular content and it is protected. The second one is the "*unprivileged partition*" (where it is used the LFU algorithm) in which the replacement of a popular content item is required and the content is then send to the privileged partition.

Count on the proposed scenario, when a user, connected to the wireless network, request access to a web place, the CDA of the CDS "*examines*" the occasion, depending on the request, and then provides the requested access through the predetermined in each case a server. Thus, after the inspection to be carried out by

CDA, if the user chooses a secured network space to access the connection will be routed through the CSe, while if the user chooses a not secured network space to access the connection will be routed through the ESe. In many case, due to the major characteristics of each of Cloud and Edge technologies, the requested access routed by default through a specific server, depending on how a similar connection was set up in the past.

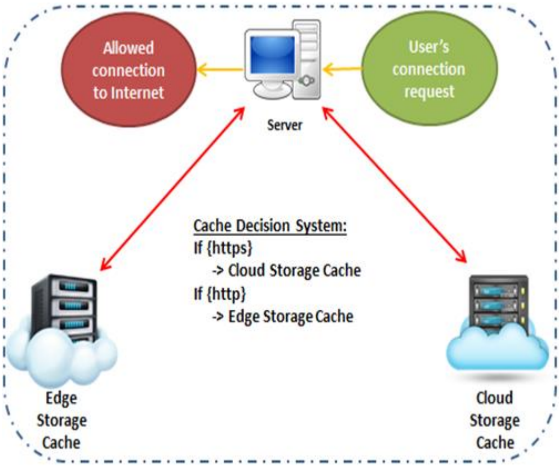


Figure 11.3: Proposed Cache Decision System (CDS).

11.3.2. Algorithm Approach of CDS

Except the algorithms, there are also caching frameworks. Some of them are the *EHCache* and the *WhirlyCache*. Both frameworks support the LRU and LFU algorithms and that means, both support LFRU the same. Because of this, these frameworks fit to the proposed caching system. After performance evaluation tests carried out by several researchers the *EHCache* framework is the one that performs better for cache miss and cache hits, and offers characteristics such as distributed functionality and monitoring statistics. When saying cache hit it means that the content-item desired after checked if available in cache storage it can be used. Even though, when saying cache mish that means that the desired content-item has not been found in the cache storage.

The CDS is the system which is responsible for the cache replacement operation. In our case study, the CDS considers if the requested content is protected (e.g. HTTPs) or if it is just a simple web page (e.g. HTTP). Figure 11.3 above demonstrates the CDS operation of the Proxy Server. In the Algorithm 1 below, the pseudocode of the proposed CDS is presented.

Algorithm 1: Cache Decision Algorithm (CDA)

```
LS= Local Server
ESe= Edge Server
CSe= Cloud Server
PS= Proxy Server
x= data-item
y= content-item HTTPs to be cached
z= content-item HTTP to be cached
for each x do
    check with PS
    if x=y then
        if x found in CSe then
            open content-item
        else
            store x in CSe with LFRU
            open content-item
    else if x=z then
        if x found in ESe then
            open content-item
        else
            store x in ESe with LFRU
            open content-item
    else
        break
    end if
end for
```

Algorithm 1 represents the *CDA* procedure operating on *CDS* of the proposed scenario, during the interaction with a user who tries access a web page. In particular, when the user requests a content-item its *PS* request evaluated with the data-item *x*. When this request received the data-item *x* evaluated and checked if it is related even to *HTTP* or *HTTPs*, in order to proceed accordingly to the next procedure. Then, depending on the case, it follows the corresponding steps. In case of $x = HTTPs$ (value of *y*), the content-item is re-evaluated if it is already exist, by a previous connection, in the *CSe* storage space, and if it is not the content-item stored with the use of *LFRU* algorithm. Respectively, in case of $x = HTTP$ (value of *z*), also, the content-item is re-evaluated if it is already exist, by a previous connection, in the *ESe* storage space, and if it is not the content-item stored with the use of *LFRU* algorithm. Consequently, in both cases with the use of *LFRU* algorithm that operates simultaneously and subsequently of the proposed algorithm the user would be able to access the requested web page. In each case, the running time of Algorithm 1 (*CDA*) is estimated less than 3 to 5 seconds, regarding the need of quick response to the client and relaying to the calculations which need to be done in any case. The *CDA* improves the communication of the proposed system providing more direct and faster response of the *LS* to each user requests a connection. *CDA* scenario leads us to leads us to a more energy efficient operation of the system, which is presented in more detail in Section 11.5 below. Moreover, regarding the complexity of *CDA* is very low due to lack of multiple variables, and the importance of the hardware infrastructure in each occasion.

11.4 Comparative Analysis

The study of previous works cites us relevant frameworks and systems proposals for assorted types of wireless networks, based in different cache types, which on several occasions supported and combined with other technologies, such as BD and CC. In this section we will make a brief study of relative works and proposals of frameworks with cache-based scenarios.

On the study conducted we singled out four (five) previous framework proposals relating to wireless networks with cache scenarios. Here, there will be a comparison between some features and the technical specifications of each proposed system.

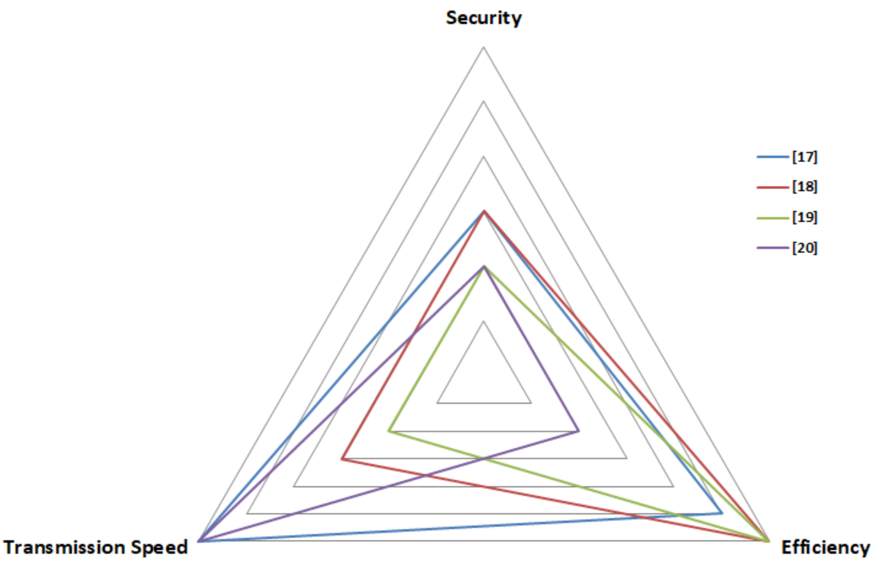


Figure 11.4: Comparative analysis of basic Framework’s features.

Relative Works	[17]	[18]	[19]	[20]
Based Network	Wireless / Wired	5G wireless network	5G wireless network	Mobile wireless network
Cache Type	Local-based Conventional Server	Cloud-based (CDH4)	Cloud-based	Local-based Conventional Server
Supported Types	all types	all types	all types	all types
Mobility	No	Yes	Yes	Yes

Table 11.1: Framework Comparative Analysis

As we can observe from Table 11.1 and Figure 11.4 most former works deal with the Efficiency of the network. On the hand, the relative works which studied in deal less with the Security of the network. Also, another feature of the wireless networks is the Transmission Speed, but only two of the four relative works deal with this feature, with the aim to provide high-speed transmission. Moreover, to offer a better analysis of Figure 11.4, we can specify that the ‘short’ cylinders describe the Low contribution and the ‘tall’ cylinders describe the High contribution. In addition to this, as we can observe from Table 11.1, the most proposed of network scenarios use Cloud-based cache storage, this could initial lead us that CC recommended for cache systems, but in addition we have the opportunity to propose a new scenario for cache systems based on EC. Furthermore, both four former works used a wireless network for their proposed scenarios, in two cases this wireless network based on 5G. Another notable feature drawn from Table 11.1 is the mobility that 3 of the 4 former works adopted into their scenario. Count on Table 11.1 and Figure 11.4 we can conclude that the proposed new scenario would be good to operate through a modern wireless network, including the feature of mobility and certainly assisted by a Cloud-based server or in general a Fog-based server.

Thus, the main purpose of these works is to provide secure and quality transmission procedure through a novel wireless network environment, such 6G, based on an integrated model for the cache system.

11.5 Experimental Results

By implementing the proposed scenario-framework in our campus, we have tested its efficiency and its security on sharing and managing our data.

We perform a number of measurements through which we can realize that we have done a good effort, but however there also many other improvements need to be done.

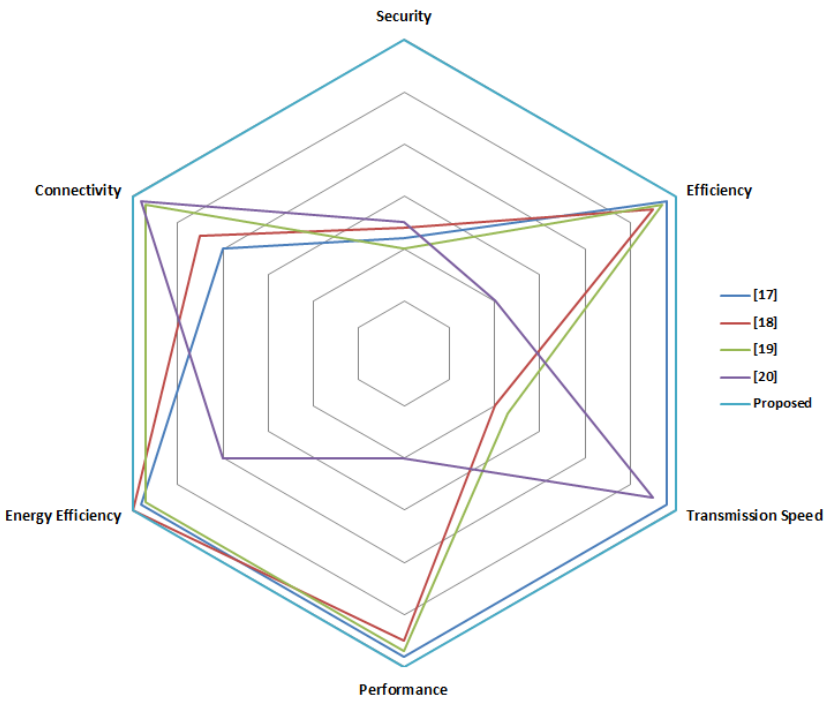


Figure 11.5: Frameworks comparison including proposed.

Figure 11.5 demonstrates the features from former works frameworks compared to our proposed framework. The features that we choose to compare are: *Security*, *Efficiency*, *Transmission Speed*, *Performance*, *Energy Efficiency* and *Connectivity*. As we can observe, our proposed scenario reaches the "*High Level*" of contribution and operation of these features than the former scenarios. We have better contribution in the Security level, which is the main objective in data, due to their value. For better understanding, we offer Table 11.2 which lists the compares features of all frameworks too in a different point of view.

Relative Works	**[17]**	**[18]**	**[19]**	**[20]**	**Proposed**
Security	Low	Low	Low	Low	High
Efficiency	High	High	High	Low	High
Transmission Speed	High	Low	Low	High	High
Performance	High	High	High	Low	High
Energy Efficiency	High	High	High	Medium	High
Connectivity	Medium	Medium	High	High	High

Table 11.2: Frameworks Comparison Including Proposed

Moreover, Table 11.2 demonstrates the selected features as they are mentioned/used in each of the previous works studied and compared here. In contrast to the previous work that has been done, we notice that our proposal is more concerned with the *Security*. While, we can observe that like our proposal so most of the former proposals (3 out of 4) deal with both *Energy Efficiency* and *Performance*.

With the aim to study further the power consumption and the efficiency of the proposed system we have adapted and implemented the following equations:

$$PC = PO * (TT + CT + RT) \qquad (1)$$

$$EE = {(DA/CS * PO)}/{PC} \qquad (2)$$

Abbreviation	Description	Abbreviation	Description
CS	Cache Speed	*PC*	Power Consumption
CT	Connection Time	*PO*	Power
DA	Data	*RT*	Reply Time
EE	Energy Efficiency	*TT*	Transmission Time

Table 11.3: Data Used

Table 11.3 demonstrates what each abbreviation represents. Moreover, it shows all the information we used in order to calculate the energy efficiency of our proposed system.

Equation (1) calculates the Power Consumption (PC) of the system by multiplying the Power used with the summary of all Time, which are Transmission Time, Connection Time and Reply Time. All time parameters affect the energy consumption of the system, as based on each time; more energy is consumed in the system.

Equation (2) calculates the Energy Efficiency (EE) of our proposed system by dividing the product of the division of Data with Cache Speed and Power, all this divided with the Power Consumption of the proposed system.

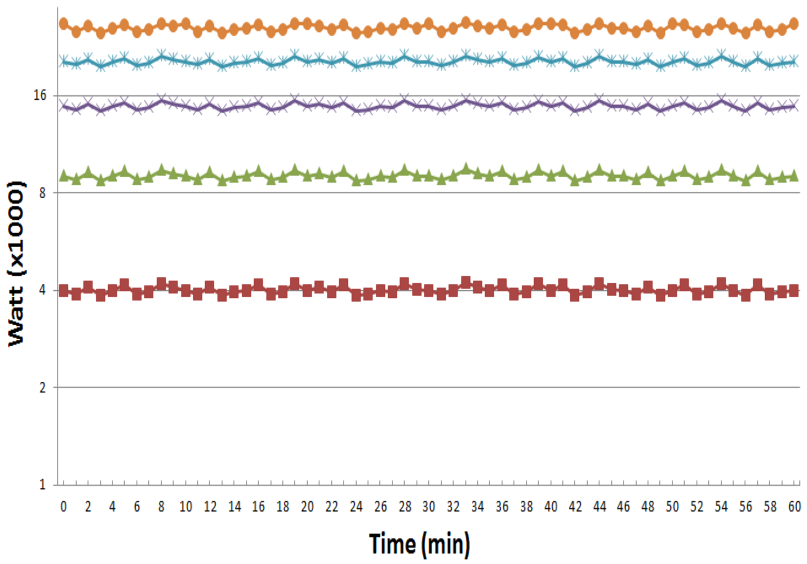


Figure 11.6: Frameworks comparison including proposed.

Initially, to be more specific, in Figure 6 the following clarifications are given: orange line represents Raichura & Padhariya work [20], light blue line represents Jin et al. work [18], purple line represents Li et al. work [19], green line represents Luo et al. work [17], and red line represents our proposed system. Figure 11.6 presents the Power Consumption (PC) of our proposed system in the duration of time, compared to other systems studied in this work. We can realize that our system could be characterized as energy efficient due to its unchanged energy consumption during its operation and the less power needed compared to the other systems.

Concluding, we can add a System of Extremely Scale Analytics tool in our proposed system in order to exploit all the data stored in our Cache Servers in order to extract useful information about the use and the reliability of the network. Something that we can deduce in the first instance, relaying on the results presented in Figure 11.6.

Systems	**Time (min)**							
	0	5	10	15	20	25	30	
Proposed	4	4	4	4	4	4	4	**Watt (x1000)**
[17]	5	5	5	5	5	5	5	
[18]	5,8	5,8	5,8	5,8	5,8	5,8	5,8	
[19]	5,5	5,5	5,5	5,5	5,5	5,5	5,5	
[20]	6,5	6,5	6,5	6,5	6,5	6,5	6,5	

Table 11.4: Power Consumption Comparison

Table 11.4 additionally represents with more information the Power Consumption of the related systems in comparison with our proposed system in numerical details. In detailed, Table 11.4 demonstrates the power consumption measured in Watts at 7 time points within the time period measured in minutes. These numbers / elements resulted from the application of equations (1) and (2) both in the previous 4 works and in our proposed scenario.

Furthermore, Figure 11.7, Figure 11.8 and Figure 11.9 describe the better Energy Efficiency offered in an academic framework implementing our proposed scenario. As we can observe our proposed scenario needs less energy power than the others through the time of its use. This can offer to the academic society a better option of energy consumption and offer a more environmental friendly framework.

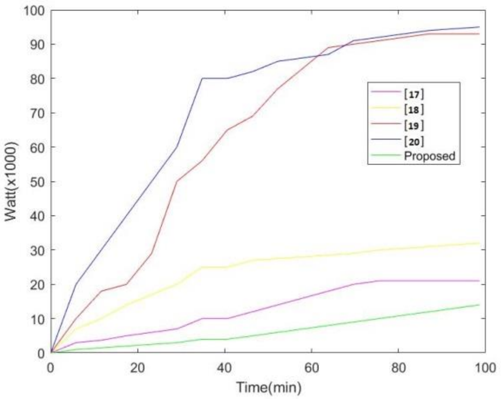


Figure 11.7: Energy Efficiency Comparison (a).

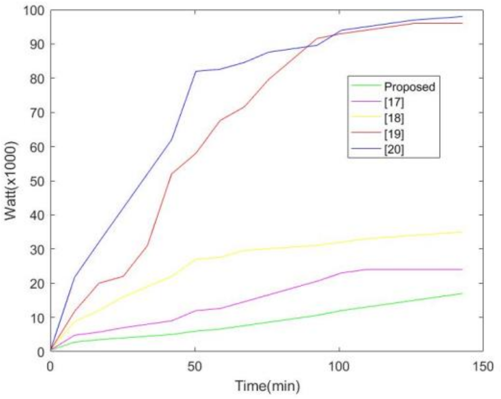


Figure 11.8: Energy Efficiency Comparison (b).

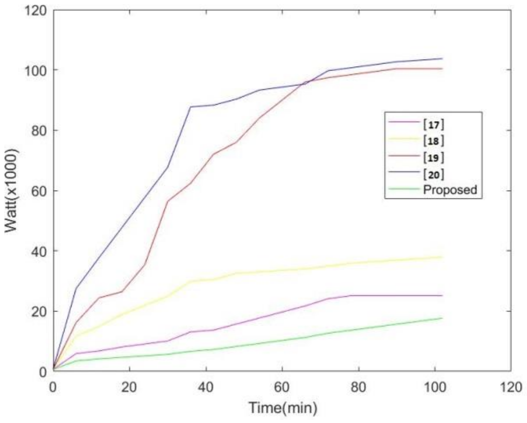


Figure 11.9: Energy Efficiency Comparison (c).

Furthermore, in Figures 11.7, 11.8 and 11.9 the following clarifications are given: blue line represents Raichura & Padhariya work [20], yellow line represents Jin et al. work [18], red line represents Li et al. work [19], purple line represents Luo et al. work [17], and green line represents our proposed scenario.

11.6 Chapter Summary

With this work we aim to offer a better scenario on a Cache Decision System of a Smart Building established on a University campus through a wireless network. Through this wireless network we can combine the functions of IoT, CC, EC and BD which play a vital role in telecommunications field. Regarding the telecommunications sector, the wireless network of the intelligent-smart building was based on 6G technology, with all the benefits this technology offers. Thus, as a main purpose, this work proposes a novel CDS through a 6G wireless network, which will offer to the users, safer and efficient environment for browsing the internet, sharing and managing large-scale data in the fog. The proposed CDS consisted of two types of servers, one CSe and one ESe. In order to come up with our proposal, we have studied related caching system scenarios from former works, which are listed and presented in this paper.

As future work, it is planned to improve our proposed system and investigate for an even better CDA for the CDS in which all necessary configurations will be taken into consideration, such as every services that will be provided by the Intelligent Building to the users that have access on it. Last but not least, this research could be a start point for better and more efficient wireless networking scenario, for managing and sharing Big Data on a Smart Building. Furthermore, as an extension of the proposed scenario, it is planned to be tested on a Smart Hospital, due to the meaning of the data used and transferred on a hospital, which are needed to be secured and immediately accessible. Thus, as case study for the future, we can implement an Extremely Scale Analytics System (ESAS), relaying on the experimental results of this work. The ESAS would be installed in the central server of proposed CDS, and so it will take advantage of the system's efficient operation settled in the SB.

11.7 Chapter References

[8] Q. Qi, X. Chen. C. Zhong, Z. Zhang, "Integration of Energy, Computation and Communication in 6G Cellular Internet of Things", IEEE Internet of Things Journal, vol. 24, issue: 6, pp. 1333 - 1337, June 2020. [DOI: 10.1109/LCOMM.2020.2982151]

[9] M. Shen, X. Tang, L. Zhu, X. Du, M. Guizani, "Privacy-Preserving Support Vector Machine Training Over Blockchain-Based Encrypted IoT Data in Smart Cities", IEEE Internet of Things Journal, vol. 6, issue: 5, pp. 7702-7712, October 2019. [DOI: 10.1109/JIOT.2019.2901840]

[12] D. Liu, Z. Yan, W. Ding, M. Atiquzzaman, "A Survey on Secure Data Analytics in Edge Computing", IEEE Internet of Things Journal, vol. 6, issue: 3, pp. 4946 – 4967, June 2019. [DOI: 10.1109/JIOT.2019.2897619]

[15] C. Esposito, M. Ficco, A. Castiglione, F. Palmieri, H. Lu, “Loss-Tolerant Event Communications Within Industrial Internet of Things by Leveraging on Game Theoretic Intelligence”, IEEE Internet of Things Journal, vol. 5, issue: 3, pp. 1679-1689, June 2018. [DOI: 10.1109/JIOT.2017.2782264]

[17] Z. Luo, M. L. Wang, Z. Lin, L. Huang, X. Du, M. Guizani, “Energy-Efficient Caching for Mobile Edge Computing in 5G Networks”, Applied Sciences, vol. 7, issue 6, pp. 1-13, June 2017.

[18] H. Jin, D. Xu, C. Zhao, D. Liang, “Information-centric mobile caching network frameworks and caching optimization: a survey”, CrossMark, Springer, EURASIP Journal on Wireless Communications and Networking, pp. 1-32, February 2017.

[19] T. Li, T. Jung, Z. Qiu, H. Li, L. Cao, Y. Wang, “Scalable Privacy-Preserving Participant Selection for Mobile Crowdsensing Systems: Participant Grouping and Secure Group Bidding”, IEEE Transactions on Network Science and Engineering, vol: PP, issue: 99, January 2018.

[20] K. Raichura, N. Padhariya, “BigCache: a cashe-based BigData management in mobile networks”, International Journal in Mobile Communications, vol. 15, no. 1, pp. 49-68, 2017.

[22] Z. Su, Q. Xu, Q. Qi. “Big Data in Mobile Social Networks: A QoE-Oriented Framework”, IEEE Network, February 2016.

[23] M. Meoni, R. Perego, N. Tonellotto, “Dataset Popularity Prediction fo Caching of CMS Big Data”, Springer, Journal of Grid Computing, pp. 1-18, February 2018.

[24] Y. Djemaiel, S. Berrahal, N. Boudriga, “A Novel Graph-Based Approach for the Management of Health Data on Cloud-Based WSANs”, Springer, Journal of Grid Computing, pp. 1-28, March 2018.

[25] A. Rafique, D. Van Landuyt, W. Joosen, “PERSIST: Policy-Based Data Management Middleware for Multi-Tenant SaaS Leveraging Federated Cloud Storage”, Springer, Journal of Grid Computing, pp. 1-30, March 2018.

[26] J. Fang, A. Ma, “"IoT Application Modules Placement and Dynamic Task Processing in Edge-Cloud Computing”, IEEE Internet of Things Journal, Early Access, pp.1-1, July 2020. [DOI: 10.1109/JIOT.2020.3007751]

[27] K. E. Psannis, C. Stergiou, B. B. Gupta, “Advanced Media-based Smart Big Data on Intelligent Cloud Systems”, IEEE Transaction on Sustainable Computing, vol. 4, Issue: 1, pp. 77-87, January-March 2019.

[28] C. Stergiou, K. E. Psannis, B. Gupta, Y. Ishibashi, “Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT”, Elsevier, Sustainable Computing, Informatics and Systems, vol. 19, pp. 174-184, September 2018.

[29] S. P. Ahuja, N. Wheeler, “Architecture of Fog-Enabled and Cloud-Enhanced Internet of Things Applications”, IGI Global, International Journal of Cloud Applications and Computing (IJCAC), vol. 10, issue. 1, pp. 1-10, January 2020. [DOI: 10.4018/IJCAC.2020010101]

[30] J. A. Jeba, S. Roy, M. Or Rashid, S. T. Atik, Md Whaiduzzaman, “Towards Green Cloud Computing an Algorithmic Approach for Energy Minimization in Cloud Data Centers”, ACM, International Journal of Cloud Applications and Computing (IJCAC), vo. 9, No. 1, pp. 59-81, January 2019. [DOI: 10.4018/IJCAC.2019010105]

Chapter 12

Advanced Media-based Smart Big Data on Intelligent Cloud Systems

K. E. Psannis, C. Stergiou, B. B. Gupta, "Advanced Media-based Smart Big Data on Intelligent Cloud Systems", IEEE Transaction on Sustainable Computing, vol. 4, Issue: 1, pp. 77-87, January-March 2019.

12.1 Introduction

Today's advanced media technology preaches an enthralling time that will enormously bear on daily life. Moreover, the rapid rise of wireless communications and networking will ultimately bring advanced media to our lives anytime, anywhere, and on any device. According to the National Institute of Standards and Technology (NIST), Cloud Computing (CC) is a scheme for enabling convenient, on-demand network access to a shared pool of configurable computing pores (for example networks, applications, storage, servers, and services) which could be promptly foresighted and delivered with minimal management effort or service provider interaction. This paper proposed an efficient algorithm for advanced scalable Media-based Smart Big Data (3D, Ultra HD) on Intelligent Cloud Computing systems. The proposed encoding algorithm outperforms the conventional HEVC standard which is demonstrated by the performance evaluations. To ratify the proposed approach addition, a relative study has been carried out. The proposed method could be used and integrated into HEVC, as a Smart Big Data, without violating the standard.

12.2 Related Work

During the last years, several techniques for HEVC-media have been contrived [19-45].

In [19] researchers proposed an innovative perception-based quantization with the aim to remove nonvisible information in high dynamic range (HDR) color pixels by exploiting luminance masking so that the performance of the HEVC standard is meliorated for HDR substance. Specifically, profile scaling count on a tone-mapping curve computed for each HDR frame is introduced. The suggested method in [19] has been integrated the HEVC relation model for the HEVC range extensions (HM-Rext). The performance of this proposed process was assessed by numbering the bitrate depletion against the HM-Rext. At the end, the results compared with HEVC at the same quality indicate that the recommended method accomplishes important bitrate savings, up to 42.2%, with an average of 12.8%. The [20] presents a computationally scalable algorithm and its hardware architecture able to support intra encoding up to 2160p@30 frames/s resolution. More specifically, the scalability permits a tradeoff between the throughput and the compression efficiency. Respectively, the encoder is capable of checking a variable number of nominee modes. Also, for the processing of the predictions that generated from the reconstructed samples uses the shared hardware resources. Thus, with the aim to support intra 4×4 modes for the 2160p@30 frames/s resolution, the encoder incorporates a separate reconstruction loop. With the aim to decrease the complexity of the quantization process in HEVC with RDOQ, [21] investigated two schemes (the RDOQ bypass decision and the simplified level adjustment. Additionally, the simplified level adaptation method only estimates the difference in rate–distortion costs among the candidate quantization levels with the aim to enable the encoder for selecting an optimal quantization level at a much reduced computational cost. Moreover, the proposed simplified level adjustment scheme is designed so that it could be implemented in lookup tables.

Also, in [22] there is an introduction of two ways to reduce the complexity of the inter/intra-mode search process of the to-be-encoded blocks in the dependent texture views (DVts) of 3D-HEVC. Also, there is a proposal of a hybrid complexity reduction scheme that utilizes the two-mode prediction approaches, which are the motion information of the base texture view (BVt), and the rate deformation cost of the already encoded blocks in the BVt and DVt. Concluding, the proposed evaluations confirmed that hybrid complexity reduction scheme reduces the 3D-HEVC codec complexity by 67.70% on average for the DVt compared with the 3D-HEVC encoder. In the other hand, the state-of-the-art method reduces complexity by 25.74% on average. In [23] for the HEVC standard and its multiplayer extensions, including SHVC and MV-HEVC extensions, there is a proposal of asoftware parallel decoder architecture. More specifically, the suggested multilayer HEVC decoder is collateral friendly and patronages both wavefront parallelism to simultaneously procedure adjacent rows of the frame and frame-based parallelism to decode a set of temporal and spacial frames in parallel. Then, the authors of [24] with the aim to decrease the encoding complexity for HEVC proposed an all-zero block detection scheme prior to DCT. Also, in [24] a modern AZB detection scheme is proposed for the case that Hadamard transform is used as a deformation metric for RDO in HEVC. At the end, the experimental results of [24] show that the proposed scheme descries 87.79% of actual AZBs with 2.87% false alarm rate in average, outperforming the state-of-the-art method. The [25] delineates an extension of HEVC standard for coding of multi-view video and depth data. The proposed approximation offers 50% bit rate savings in contrast with HEVC simulcast and 20% in comparison with a straightforward multi-view extension of HEVC without the newly developed coding tools demonstrated in [25] by the objective and subjective results presented. Moreover, a fast CU size decision algorithm for HM was proposed by the authors of [25].

Moreover, in [26] the authors can determine CU depth range and skip some specific depth levels rarely used in the former frame and neighboring CUs. Finally, the proposed algorithm could particularly decrease computational complexity while maintaining nearly the same RD performance as the original HEVC encoder as these presented in [26] by the experimental results. A comparison between the compression capability of sundry generations of video coding standards and the means of peak signal-to-noise ratio (PSNR) and substructure testing results provided in [27]. Furthermore, the authors of [27] applied a consolidated approach to the analysis of designs, including H.262/MPEG-2 Video, H.263, MPEG-4 Visual, H.264/MPEG-4 AVC, and HEVC. Concluding, equivalent substructure reproduction quality as encoders, which conforms to H.264/MPEG-4 AVC when using approximately 50% less bit rate on average, could be reached by the results of subjective tests for WVGA and HD sequences point out that HEVC encoders. Also, in [28] there is a consideration of a classification of motion activity. Concluding this paper, the experimental results exhibit that the encoding complexity could be decreased by up to

38% on average in the random access main profile configuration with only a small bit-rate growth and a peak signal to noise ratio (PSNR) decrement, compared to HEVC test model (HM) 7.0 reference software. An algorithm that provides the possibility of applying for both CU and PU parts is proposed by the authors of [29]. Moreover, there is a proposal of a CU splitting algorithm based on the rate–distortion cost of CU about the parent and current levels to terminate the CU decision early, with the aim to reduce the computational complexity. Also, in terms of PU, the authors of [29] develop fast PU decision counted on spatio-temporal and depth correlation for PU level. Furthermore, the [30] with the aim to alleviate the encoder computation load offers a modern method to decrease the candidates in RDO process. Moreover, the proposed scheme offers 20% and 28% time savings in intra high efficiency and low complexity cases on average compared to the default encoding scheme in HM 1.0 with almost the same coding efficiency demonstrated by the experimental results of [30].

The [31] proposed a HEVC standard which is fulfilled its target to reach more than 50% improvement in video compression over the existing H.264 Advanced Video Coding standard. Concluding the work of [31], the experimental results demonstrate that for low-delay wireless video communications, the HEVC codec is more effective compare to the previous H.264 codec and shows better overall performance. In [32] described the complexity-related aspects that were considered in the standardization process. Moreover, a clue of where HEVC may be more complex than its predecessors and where it may be simpler was given by profiling of reference software and optimized software. The [33] offers an overview of the intra coding techniques in the HEVC standard being produced by the Joint Collaborative Team on Video Coding (JCT-VC). Also, the [33] discusses the design principles applied during the development of the new intra coding methods and additionally analyzes the compression performance of the individual tools. The [34] proposes a reusable design for the merging process used in 3D-HEVC, which could importantly decrease the implementation complexity by eliminating duplicated module redundancies. Finally, this proposed method of [34] has been adopted as a regulative coding tool in the 3D-HEVC international standard.

The contribution that the authors of [35] made fills the gap in enabling compliant and real-time networked HEVC visual applications. Also, it is taken farther by evaluating the transmission of 4k UHDTV HEVC-coded content in a typical wireless environment using both computers and mobile devices, with the aim to consider well-known factors like prevention, intervention and other unseen factors that affect the network performance and video quality. The [36] illustrates extensions to the HEVC standard which are active fields of current development in the relevant international standardization committees. The design for the extensions proposed in [36] deputizes the latest state of the art for video coding and its applications. The [37] introduces the SHVC standard technology and analyzes the performance of inter-layer prediction as an important characteristic. Finally, the authors of [37] proposed a gradient based fast decision algorithm, as a result of the fact that every coding unit with different sizes is traversed in both procedures makes it very time-consuming, with the aim to reduce the computational complexity of HEVC. As compared to the default encoding scheme in HEVC test model HM 4.0, experimental results of [37] demonstrate that the fast intra mode decision scheme offers almost 20% time savings in all intra low complexity cases on average with negligible loss of coding efficiency.

In recent years several studies for BD technologies have been devised [39-45]. The authors of [39] introduce a multi-objective approach using genetic algorithms. The goal of this is to minimize two objectives, the execution time, and the budget of each node executing the task in the cloud. The contribution of [39] is to propose an innovative adaptive model to communicate with the task scheduler of resource management. The proposed model periodically queries for resource consumption data and uses to calculate how the resources should be allocated to each task. Through this work, the authors believe that the proposed solution is timely and innovative as it provides a robust resource management where users can perform better scheduling for BD processing in a seamless manner. The [40] makes three contributions to the Special Issue's theme of enhancing organizational resource management. One is to establish an archetype business process for BD initiatives. The second contribution directs attention to creating a dynamic capability with BD initiatives. The third identifies drawbacks of resource-based theory (RBT) and it's underpinning assumptions in the context of BD. Moreover, in [40] there is a discussion about the lessons learnt and draws out implications for practice and business research. Also, this work’s intellectual and practical contributions are count on an in-depth case study of the European ICT Poles of Excellence (EIPE) BD initiative and evidence from the extant literature.

A literature review of BD and its related technologies, like CC and Hadoop, is presented in [41]. Also, the [41] focuses on the five phases of the value chain of BD technology and as a result examines the several representative applications of BD technology. Furthermore, in [42] the important concepts of BD technology are highlighted and also there is a discussion about the various aspects of BD. Furthermore, the authors of [42] define what BD is, and discuss the various parameters of its definition. Finally, in [42] there is a look at the process involved in the data processing and then reviewing the security aspects of BD and as a result, it proposes a new system for security of BD. Moreover, a definition and description of BI&A 1.0, BI&A 2.0 and BI&A 3.0 in terms of their key characteristics and capabilities are presented in [43]. Also, there is a report of a biometric study of critical BI&A publications, Researchers and research topics based on more than a decade of related academic and industry publications are presented in [43]. Additionally, an offer of six provocation with the aim to spark conversations about the issue of BD technology is shown in [44]. These provocations are the cultural, technological, and scholarly phenomena that rests on the interplay of technology, analysis, and mythology that dares extensive utopian and dystopian rhetoric. Finally, a multi-stakeholder approach for developing a suitable privacy regulation in the age of BD is presented in [45]. This argument is developed in five steps: 1) A review of the current academic debate on privacy regulation. 2) An argument that the framework for developing a suitable privacy regulation should not only focus on formal and procedural but also include some essential aspects to guard users and promote socially beneficial BD applications. 3) An examination of how the process leading to an appropriate regulation might be organized. 4) A discussion of the potential structure of a privacy organization that might conduct multistakeholder-dialogues as a preliminary step. 5) A discussion of their findings and suggestions.

In the sector of CC, there is also a majority of works available for study [46-52]. At the beginning, an exploration of the roadblocks and solutions to provide a trustworthy cloud computing environment are presented in [46]. CC is an evolving paradigm with tremendous momentum, but its specific aspects sharpen security and privacy challenges. Then, the [47] proposes a simple data protection model where data is encrypted by the use of the AES algorithm before it is launched in the cloud, thus ensuring data confidentiality and security. The key consideration dealt in the proposal of [47] is the encryption schema to secure data by making it unintelligible for all. Regarding [47], implementing AES for security over data provides benefits of less memory consumption and less computation time as compared to other algorithms. Furthermore, a survey of the Mobile Cloud Computing (MCC) was given in [48], which has the purpose of helping general readers that have an overview of the MCC including the definition, the architecture and the applications. Also, it was presented the issues, the existing solutions and the approaches of them in [48]. Finally, [48] suggest the future research directions of the MCC technology.

The [49] provides an extensive survey of MCC research while highlighting the particular concerns in MCC. Furthermore, a taxonomy count on the key issues in the area of MCC and a discussion about the different approaches that have been taken to tackle these issues were presented in [49]. At the end, there is a conclusion with a critical analysis of challenges that have not yet fully met, and highlight directions for the future work. In [50] it was detailed the security issues that arise as a result of the very nature of cloud computing. Moreover, the [50] presents the recent solutions that were presented in its literature with the aim to counter the security issues. Also, a brief view of security vulnerabilities in the MCC is highlighted. Finally, it presented the discussion on the open issues and future research directions. At the end, regarding the CC technology, the [51] offers a study on CC and suitable algorithms for load balancing such as ground robin scheduling, MapReduce algorithm, ACO and honeybee. Also, it was given a comparison between those algorithms on different properties of them. According to the [51], the ACO is the better load balancing algorithm compared to other algorithms. Concluding the first part of the related review, there is a study about Big Data technology. In [52] initially, there is an investigation of the importance of BD in modern life, and in terms of the economy, and also discussed the challenges that arise from Big Data utilization. Moreover, in

[52] the potential of the powerful combination of BD and Computational intelligence is explored and a number of areas where novel applications in real world problems can be developed by utilizing these powerful tools and technologies is identified. To solve these problems, the authors of [52] presented an innovative data modelling methodology which introduces a novel biologically inspired universal generative modelling approximation called Hierarchical Spatial-Temporal State Machine (HSTSM). Finally, there is a discussion of various implications of policy, protection, valuation and commercialization related to BD, its applications and its deployment.

12.3 Integration of 3D Ultra HEVC Smart Big Data on Intelligent Clouds

Among all types of data in the Cloud storage, video has occupied a significant part because of the explosive video sharing on social networks and video-on-demand services for movies, TV programs, etc. For example, YouTube has claimed in 2015 that there are 400 hours of video uploaded to YouTube every minute. Moreover, to support users with various bandwidth requirements and device resolutions and full interactive playback in video streaming, usually, multiple versions at different bitrates, resolutions and frame rates are generated for each video, which is called simulcast in video streaming [31]. An alternative to satisfy these adaptive streaming requirements but with less storage is the scalable video coding (SVC), e.g. H.264/SVC and HEVC/SHVC, where a video is coded into one base layer (BL) and several enhancement layers (EL) [32] [37] [53].

A new challenge is created when every user has to manage and process big quantities of data everywhere and every time. This challenge is to use Smart BD in Intelligent CC, and though this new general challenge, other challenges arise. One main issue and big challenge for the use of BD in Intelligent Cloud environments is the transfer and the use of High Quality Video. High quality video is a new type of video coding that grows through the recent years. Some major types of High Quality Video coding are the 3D Ultra HD Video and the 3D HEVC. The main challenge of this work is to try to transfer and to use those types of videos as Smart BD through Intelligent Cloud environments.

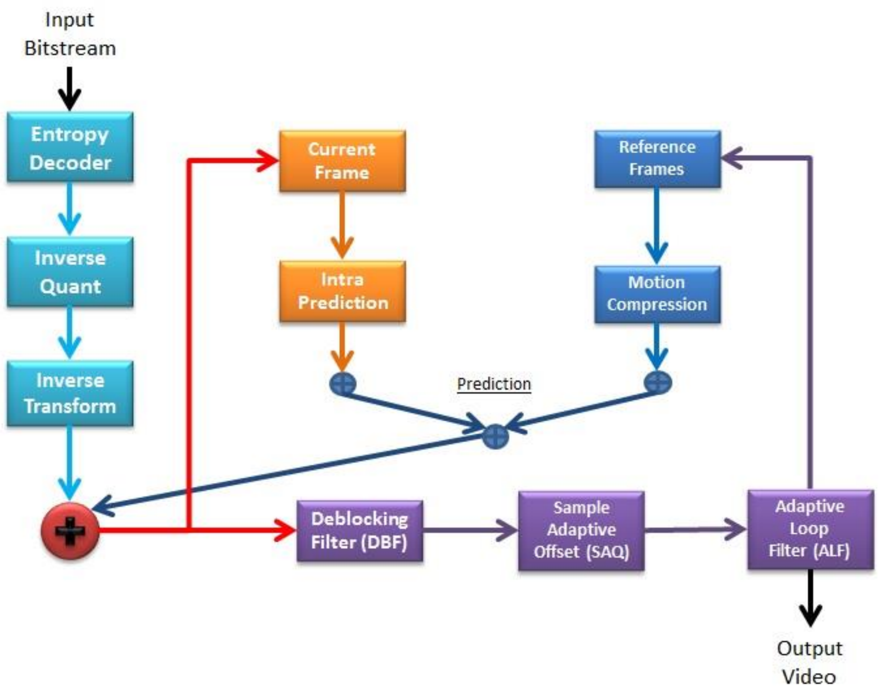


Figure 12.2: Block Diagram of HEVC Decoder.

Figure 12.2 displays the operating procedure of the HEVC Decoder. We can follow the Input Bit-Stream from its entry to the decoder to its exit as an output video. The whole procedure of the HEVC decoder could be separated in four operating sub-processes, which are singled out in Figure 12.2 by their different color.

As already mentioned above, High-Efficiency Video Coding, or better publicly recognized as HEVC, is a video compression standard, one of the several potential successors to the widely used AVC (H.264 or MPEG-4 Part 10). It provides about two times more the data compression ratio at the same level of video quality. More specifically, the 3D-HEVC extensions for 3D video were completed in early 2015. Farther extensions remain in development for completion in early 2016, covering video containing rendered graphics, text, or animation as well as camera-captured video scenes. The 3D-HEVC is the third version of HEVC coding which was released on April 29, 2015. This third version adds the 3D main profile in HEVC/H.265 coding. The 3D main profile allows for the base layer which conforms to the main profile of HEVC [54]. The 3D-HEVC offers increased coding efficiency by the joint coding of texture and depth for advanced 3D displays. Through experimental analysis [55] it was shown that the 3D-HEVC is capable of achieving the same subjective video quality as the H.264/MPEG-4 AVC High Profile while requiring on average only about 50% of the bit-rate.

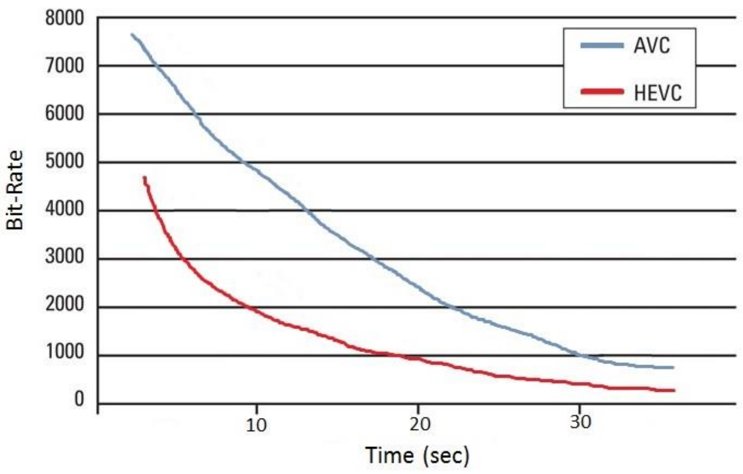


Figure 12.3: AVC/H.264 vs. HEVC/H.265.

Figure 12.3 shows the quantity of data that compressed, or otherwise how the bit-rate decreased through time compering the AVC/H.264 video coding to HEVC/H.265 video coding.

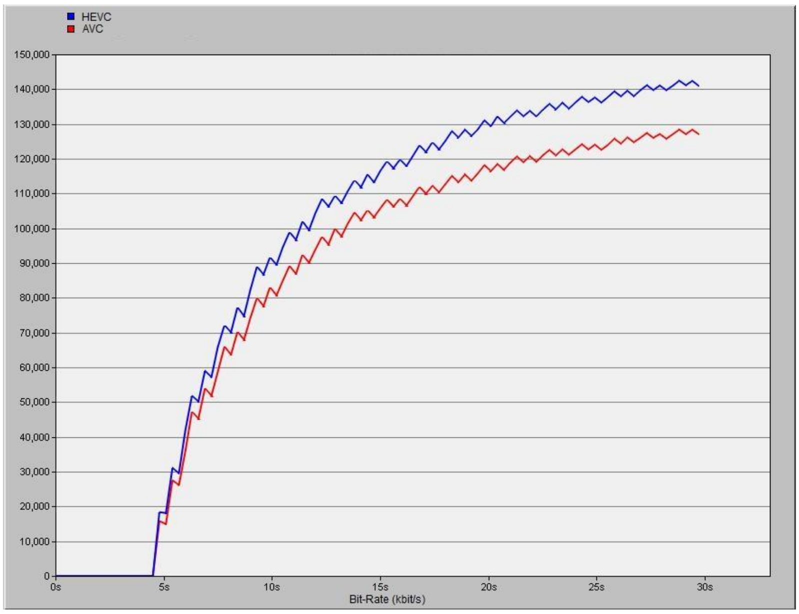


Figure 12.4: HEVC/H.265 to AVC/H.264 comparison.

Figure 12.4 reveals that for very low bit-rates of HEVC and AVC provided almost the same quality, but pointing out that HEVC is much quicker. Figure 12.3 shows the amount of bit-rate of HEV and AVC through the time.

Furthermore, another recent type of video coding which also is a base of the updated version of HEVC is Ultra High Definition or better known as Ultra HD. Ultra HD in our days includes video types of 4K UHD in 2160p resolution and 8K UHD in 4320p resolution. Those two types of video formats were first proposed by the *NHK Science & Technology Research Laboratories* and later defined and approved by the

International Telecommunication Union, better known as ITU [56] [57] [58] [59].

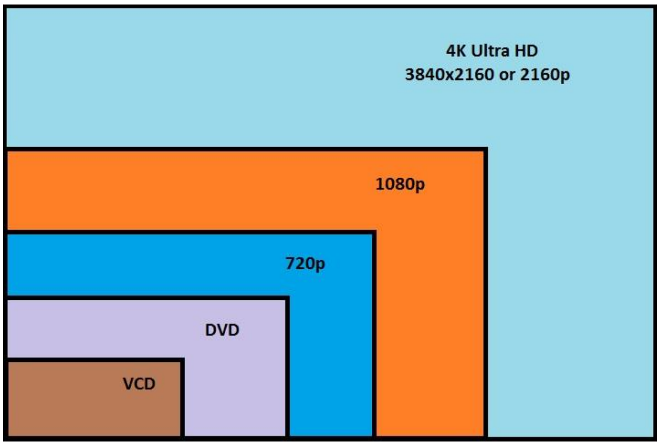


Figure 12.5: Generation of video resolution.

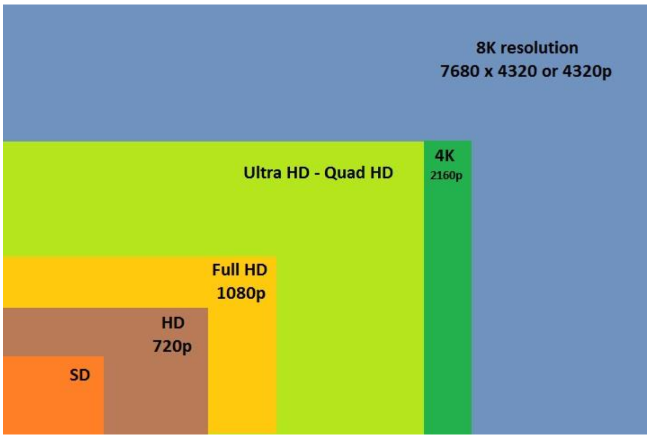


Figure 12.6: Comparison of Resolutions.

Figure 12.5 and Figure 12.6 represents the differences between the resolutions of video codecs and their dimensions. More specifically, Figure 12.5 shows the big impact of 4K Ultra HD instead of the previous resolutions and the Figure 12.6 shows the effect in the dimension of the screen derived from the 8K Ultra HD and how bigger it is, compared to the previous resolutions. Moreover, Figure 5 and Figure 6 illustrates the generation of Display Resolutions by comparing the Display Analysis of each Resolution's size.

4K and 8K resolutions that were introduced by Ultra HD can also be defined as UHDTV [60] [61]. More specifically, 4K UHDTV or 2160p consists of 3840 pixels in wide view and by 2160 pixels in tall (8.29 MegaPixels). This is four times as many pixels as the Full HD which is consisted by 1920x1080 pixels (2.07 MegaPixels). Also, the 8K UHDTV or 4320p consists of 7680 pixels in wide view and 4320 pixels in tall (33.18 MegaPixels). This video coding brings the view closer to the detail level of 15/70mm IMAX. As an evolution of this video coding on August 22, 2012, the electronics company LG [62] released the first 3D UHDTV that supports the 4K system. This came as a follow-up to the SONY's released 4K 3D Projector with model name VPL-VW1000ES [63] [64] on May 31, 2012. The new type of video coded videos related on UHD, HEVC and 3D sized by a lot of bytes.

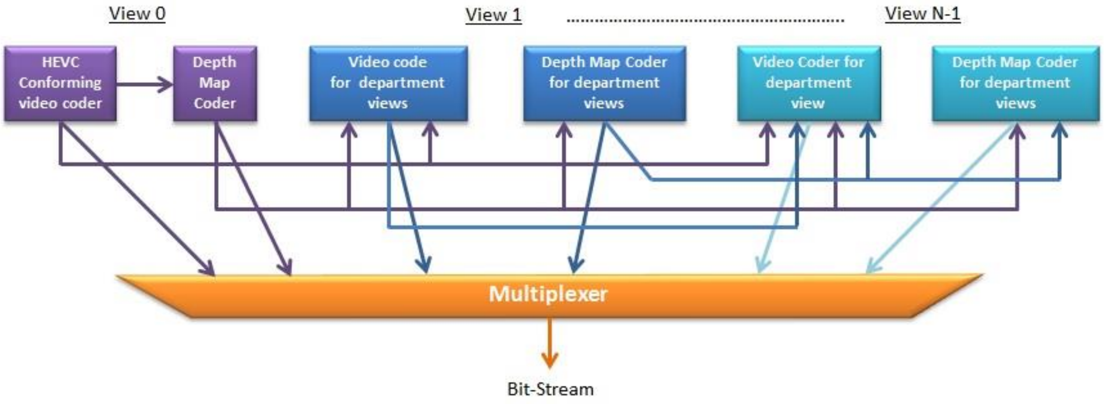

View 0
View 1
View N-1
HEVC Conforming video coder
Depth Map Coder
Video code for department views
Depth Map Coder for department views
Video Coder for department view
Depth Map Coder for department views
Multiplexer
Bit-Stream

Figure 12.7: Inter-View & Inter-Component Prediction in Basic Encoder Structure of 3D-HEVC.

Figure 12.7 shows the basic structure of a 3D video encoder of HEVC. 3D HEVC was developed for depth-enhanced 3D video formats, ranging from Conventional Stereo Video to Multi-View Video plus a depth consisting of two or more views and associated per-pixel depth data components.

Regarding the definition of BD, the data which set by a large amount of bites, can be defined as Big Data. The major issue of this work is to manage a large amount of data by 3D HEVC of 3D UHD videos, which is Big Data, through the Cloud environments. Thus, with the aim to allow every user to manage and process big amounts of data everywhere and every time a new challenge in the sector of telecommunications was created. The CC technology eliminates the need to maintain expensive computing hardware and software [65] [66] [67]. The CC resources and techniques could be influenced to address the conventional problems associated with fault tolerance and low performance causing bottlenecks to the use of BD technology [68] [69] [70]. The usage of BD offers the specific opportunity to reach an appropriate competitive strategic advantage provided to the users to use the right mix of BD analytics to discover relationships and patterns that could not be discovered otherwise [71]. Also, regarding the related review CC and BD are complementary to each other and some of the Big Data problems can be resolved with the CC techniques and solutions.

Big Data Features	***Volume***	***Velocity***	***Variety***
Cloud Computing Features			
Storage over Internet		**X**	
Service over Internet	**X**		**X**
Applications over Internet	**X**	**X**	**X**
Energy Efficiency	**X**	**X**	
Computational Capable		X	X

Table 12.2: Big Data Features Contribution in CC Features

Table 12.2 exhibits the key characteristic of the two technologies which have been studied and used with the aim to use them for the experimental proposal. Based on the study conducted, the key characteristic of BD technology which contributes more the characteristics of CC technology is Velocity. Velocity contributes four from the five key characteristics of CC. Also, another thing that we can observe from Table 3 is that the characteristic Applications over Internet contributed from all the key features of BD.

12.4 Adaptive 4K, 8K, 3D Media Delivery

Count on the CfP for 3D video technology that presented on [64], it specified that there are two test categories for the 3D video technology. Those two categories are the AVC-compatible and the HEVC-compatible/unconstrained. The proposed algorithm that introduced in this work is based on the model that presented on [72].

Test Sequence	**3-View Test Scenario AVC - BitRate**				**3-View Test Scenario HEVC - BitRate**			
	R1	**R2**	**R3**	**R4**	**R1**	**R2**	**R3**	**R4**
Ponzan_Hall2	170	280	440	740	220	320	490	780
Poznan_Street	370	780	1140	1900	420	720	1190	1960
Undo_Dancer	380	740	1160	1890	440	790	1210	2020
GT_Fly	290	580	980	1450	350	610	1090	1610
Kendo	240	400	630	1000	290	440	680	1050
Balloons	270	440	740	1160	310	490	780	1210
Lovebird1	220	390	690	1240	270	430	740	1280
Newspaper	300	420	640	870	350	460	690	910

Table 12.3: Compared Table of the Rate-Points for 3-View Test Scenario AVC and 3-View Test Scenario HEVC.

As we can observe form Table 12.3, comparing the two video technologies we can understand that there is an improvement by the use of the 3D view in the HEVC video codec. Based on the test coherences from [68] we have done measurements for both codecs and the Bit-rate results of those measurements are demonstrated in Table 12.3.

Test Sequence	**3-View Test Scenario AVC - Bit-Rate Savings %**		**3-View Test Scenario HEVC - Bit-Rate Savings %**	
	MultiView	***SimulCast***	***MultiView***	***SimulCast***
Ponzan_Hall2	19.97	40.86	22.03	44.98
Poznan_Street	12.34	46.97	14.40	51.09
Undo_Dancer	10.43	49.09	12.49	53.21
GT_Fly	18.35	53.51	20.41	57.63
Kendo	38.24	49.17	40.30	53.29
Balloons	29.10	45.13	31.16	49.55
Lovebird1	16.02	43.01	18.08	47.13
Newspaper	20.96	39.06	23.02	43.18
Average %	20.68	45.85	22.74	50.01

Table 12.4: Compared Table of Average Bit-Rate Savings of Multiview and Simulcast Extension of AVC and HEVC.

Table 12.4 shows the average Bit-Rate Savings of the measurements that have been done of MultiView and SimulCast extension of the AVC codec and the HEVC codec. Through Table 4 we can assume that the 3D-HEVC is more improved than the AVC, regarding both extensions. The measurements count on the test coherences are taken from the [72].

This work tries to link and offer the possibilities to can access on demand the video (Smart Big Data) that are saved in an Intelligent Cloud. For this purpose, we try to find the better way to Bit-Stream the availability of 3D-quality video as Smart Big Data through an Intelligent Cloud system. Through the related review and the aforementioned measurements and results, we assumed that the better way to stream a high-quality video, as a 3D-HEVC, will be through a new method.

Relaying to the work of [72] and regarding the related review we propose the following equation:

$$BW_{i+1}^{next} = BW_{i}^{last} * [(f(RTT_i, RTT_{i-1}) + g(p_i, p_{i-1}) + h(SINR_i, SINR_{i-1})) * a] + BW_{i-1}^{last-1} \qquad (1)$$

where, α is a stable that indicates the importance of each factor, f(), g(), h() are three functions which reflecting the value change of each factor compared with the *last time window*, p is representing the *packet loss rate*, RTT is representing the *Round Trip Time*, SINR is representing the *signal to interference and noise ratio*, i represents the *sequence number* of the current time window, BW_{i}^{last} represents the *bandwidth of the last time window*, represents the bandwidth of the last time window-1, and BW_{i+1}^{next} represents the *bandwidth of the next time window*, which is the time we need.

Based on the proposed equation (1) we propose also an algorithm that implements it. Additionally, with equation (1) in the proposed algorithms, we also use the $\mathrm{T_{win}}$ which represents the time window, BL which represents Base Layer and EL which represents Enhancement Layer. This algorithm is showed in the following.

Algorithm 1

```
i=0
$BW_1=R_{BL}$
Transmit $BL_1$
Monitor $BW_1^{last}$
**repeat**
Sleep for $T_{win}$
Obtain $p_i$, $RTT_i$, $SINR_i$, and all the information we need
  from the client's report
Predict $BW_{i+1}^{next}$ (or $BW_{i+1}^{next}=BW_i^{last}$) (or $BW_{i+1}^{next}=BW_i^{last}+BW_{i-1}^{last-1}$)
$\alpha$=0
$BW_{EL}$=0
**repeat**
        $\alpha$++
        if $\alpha$>= $i$
      **break**
        $BW_{EL}=BW_{EL}^{\alpha}+R_{EL}^{\alpha}$
**until** $BW_{EL}$>= $BW_{i+1}^{next}$
Transmit $BL_i$and $EL_{i+1}^1, EL_{i+1}^2, ..., EL_{i+1}^{a-1}$
Monitor $BW_{i+1}^{last}$
$i$++
**until** All video segments are transmitted
```

Algorithm 1 represents the procedure implementing the equation (1) which consists due to our study the better way to stream a High-Quality Video, such as 3D-HEVC. Thus, in Algorithm 1 we test the data from *Last Time Window* compared to the functions which reflect the value change of each factor of the equation (1). In a limited number of loops, and as long as the time is "sleeping", the algorithm "calls" the value of the *Packet Loss Rate* of each loop along with *Round Trip Time* and the *Signal to Interface and Noise Ratio* of each loop, carried out from the client which streams the video. Then, the algorithmic mapping of equation (1) calculates the better way to stream the video by choosing a calculation scenario relaying on the type of the streaming video. Inside a loop which is repeated as long as the value of BW_{EL} is greater than or equal the value of BW_{i+1}^{next}, then the calculation takes as many times as the random variable α is greater than or equal to the value i, which is the *sequence number* of the particular transmission. After the iterative loop stops, the particular video begins to be transmitted to the client by streaming it. The running time of the Algorithm 1 is estimated less than one minute due to the need of quick response to the client and relaying to the calculations which need to be done.

12.5 Experimental Results

By establishing the algorithm above we can procced the procedure of selecting and streaming the video we demand in the type of it that we demand. This procedure is showing in Figure 12.8.

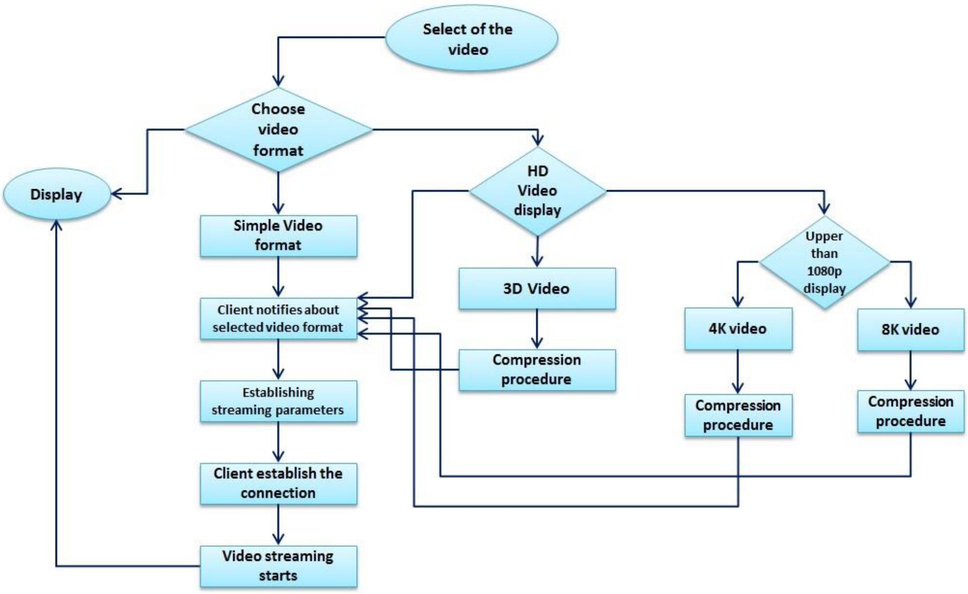


Figure 12.8: Working flow of the video streaming procedure in an Intelligent Cloud environment.

With the study and the analysis of Figure 12.8 we can observe that the client of an Intelligent Cloud environment could access the video that the client demand, and based on the available options of view, the client can choose the type of display. The waiting time of using a larger sized video format for streaming is reduced with the use of the proposed algorithm.

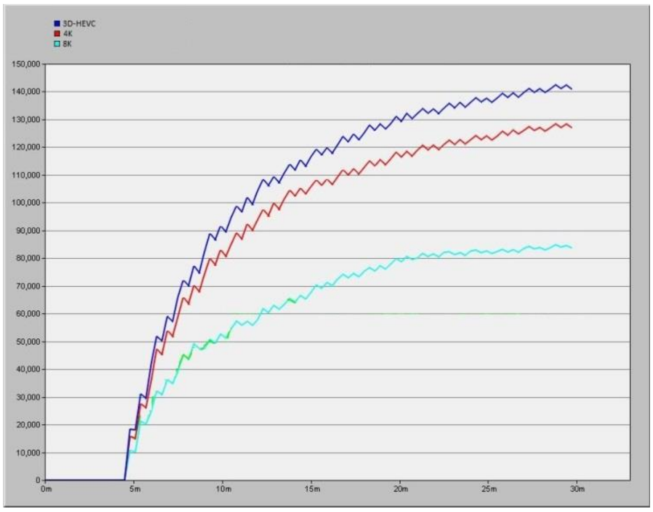


Figure 12.9: Streaming results and measurements of three video formats.

Figure 12.9 demonstrates the differences between the bit-streaming of the three video formats which were used. As we can observe there is a closer response between the 3D-HEVC and the 8K display and also not a big difference regarding 4K display. Those measurements could show that even the large data volume of the 3D-HEVC video codec, the Bit-Rate through time is not very longer than a lower type video codec as the 8K or the 4K.

12.6 Chapter Summary

In the last decades technologies like BD and Cloud became valuable for people that need information at any time in any place. Information such this can be a high quality video, e.g. a 3D-HEVC video format. In this paper, we study and survey the three aforementioned technologies in order to find their common features of their use and to propose an operation which would help the issue of streaming high quality video, as Big Data, through the cloud environments. Based on the fast growth of wireless communications and networking technologies, which are related increased in many of their features like the volume of their data in the structured and unstructured form. Also, as the technology of CC grows more options about its "on-demand" operation arise.

Thus, in this work, we proposed an efficient algorithm for advanced scalable Media-based Smart Big Data (3D, Ultra HD) on Intelligent Cloud Computing systems. With performance evaluations that have been made we demonstrate that the proposed encoding algorithm outperforms the traditional HEVC standard. By adopting this proposed method we assumed that it can be used and integrated into HEVC without violating the standard. Furthermore, by surveying the integration of BD, in general, in Cloud environments, we open new challenges in the field of this integration. This can be the sector of future research on the integration of those two technologies, and why not to have a huge improvement on their integration issues in order to have a better use of them.

12.7 Chapter References

[19] Y. Zhang,M. Naccari, D.Agrafiotis, M.Mrak, D. R. Bull, "High Dynamic Range Video Compression Exploiting Luminance Masking", IEEE Transactions on Circuits and Systems for Video Technology, vol. 5, issue. 26, pp. 950-964, May 2016.

[20] G. Pastuszak, A. Abramowski, "Algorithm and Architecture Design of the H.265/HEVC Intra Encoder", IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, issue: 1, pp. 210-222, May 2015.

[21] H. Lee, S. Yang, Y. Park, B. Jeon, "Fast Quantization Method With Simplified Rate-Distortion Optimized Quantization for an HEVC Encoder", IEEE

Transactions on Circuits and Systems for Video Technology, vol. 26, issue: 1, pp. 107-116, June 2015.

[22] H. R.Tohidypour, M. T. Pourazad, P. Nasiopoulos, "Online-Learning-Based Complexity Reduction Scheme for 3D-HEVC", IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, issue:10, pp. 1870-1883, October 2016.

[23] W. Hamidouche, M. Raulet, O. Déforges, "4K Real-Time and Parallel Software Video Decoder for Multilayer HEVC Extensions", IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, issue: 1, pp. 169-180, January 2016.

[24] B. Lee, J. Jung, M. Kim, "An All-Zero Block Detection Scheme for Low-Complexity HEVC Encoders", IEEE Transactions on Multimedia, vol. 18, issue: 7, pp. 1257-1268, July 2016.

[25] K. Müller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. H. Rhee, G. Tech, M. Winken, T. Wiegand, "3D High-Efficiency Video Coding for Multi-View Video and Depth Data", IEEE Transactions on Image Processing, vol. 22, issue: 9, pp. 3366-3378, September 2013.

[26] L. Shen,Z. Liu, X. Zhang, W. Zhao, Z. Zhang, "An Effective CU Size Decision Method for HEVC Encoders", IEEE Transactions on Multimedia, vol. 15, issue: 2, pp. 465-470, February 2013.

[27] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards-Including High Efficiency Video Coding (HEVC)", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, issue: 12, pp. 1669-1684, December 2012.

[28] J.-H. Lee,B.-G. Kim, D.-S. Jun, S.-H. Jung, J. S. Choi, "Complexity reduction algorithm for prediction unit decision process in high efficiency video coding", IET Image Processing, vol. 10, issue: 1, pp. 53-60, January 2016.

[29] J.-H. Lee, K. Goswami, B.-G. Kim, S. Jeong, J. S. Choi, “Fast Encoding Algorithm for High Efficiency Video Coding (HEVC) System Based on Spatiotemporal Correlation”, Springer, Journal of Real-Time Image Processing, vol. 12, issue: 2, pp. 407-418, August 2016.

[30] M. R. Fini, F. Zargari, "Two stage fast mode decision algorithm for intra prediction in HEVC", Springer, Multimedia Tools and Applications, vol. 75, issue: 13, pp. 7541–7558, July 2016.

[31] K. E. Psannis, "HEVC in wireless environments", Springer, J Real-Time Image Proc, vol. 12, issue: 2, pp. 509-516, August 2016.

[32] F. Bossen, B. Bross, K. Suhring, D. Flynn, "HEVC Complexity and Implementation Analysis", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, issue: 12, pp. 1685-1696, December 2012.

[33] J.Lainema, F. Bossen, W.-J. Han, J. Min, K. Ugur, "Intra Coding of the HEVC Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, issue: 12, pp. 1792-1801, December 2012.

[34] Y. S.Heo, G. Bang, G. H. Park, "Reusable HEVC Design in 3D-HEVC", Wiley, Etri Journal, vol. 38, issue: 5, pp. 818-828, October 2016.

[35] A. Adeyemi-Ejeye, M. Alreshoodi, S. D. Walker, "Implementation of 4kUHD HEVC-content transmission", Springer, Multimedia Tools and Applications, vol. 76, issue: 17, pp. 18099-18118, September 2017.

[36] G. J. Sullivan, Jill M. Boyce, Ying Chen, Jens-Rainer Ohm, C. Andrew Segall, Anthony Vetro, "Standardized Extensions of High Efficiency Video Coding (HEVC)", IEEE Journal of Selected Topics in Signal Processing, vol. 7, issue: 6, pp. 1001-1016, December 2013.

[37] C.-S. Park, B.-G. Kim, "Performance analysis of inter-layer prediction in scalable extension of HEVC (SHVC) for adaptive media service", Displays, vol. 44, pp. 27-36, September 2016.

[38] W. Jiang, H. Ma, Y. Chen, "Gradient Based Fast Mode Decision Algorithm for Intra Prediction in HEVC", in Proceedings of 2nd International Conference on Communications and Networks (CECNet), 2012, Consumer Electronics, pp. 1836-1840, Yichang, China, 21-23 April 2012.

[39] Ibrahim Abaker Targio Hashem, et al, "Schedule optimization for big data processing on cloud", in Proceedings of 2nd International Conference on Big Data Analysis and Data Mining, San Antonio, USA, 30th November-1st December 2015.

[40] A. Braganza,L. Brooks, D. Nepelski, M. Ali, R. Moro, "Resource management in big data initiatives: Processes and dynamic capabilities", Elsevier, Journal of Business Research, vol. 70, pp. 328-337, January 2017.

[41] I. Chebbi, W. Boulila Imed, R. Farah, "Big Data: Concepts, Challenges and Applications", Springer, Computational Collective Intelligence, vol. 2, p. 638–647, October 2015.

[42] R. Toshniwal, K. G. Dastidar, A. Nath, "Big Data Security Issues and Challenges", International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 2, issue: 2, pp. 15-20, February 2015.

[43] H. Chen, R. H L Chiang, V. C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact", MIS Quarterly: Management Information Systems, vol. 36, issue: 4, pp. 1165-1188, December 2012.

[44] K. Crawford, D. Boyd, "Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon", Information, Communication & Society Journal, vol. 15, issue: 5, pp. 662-679, May 2012.

[45] M. G. Will, "Privacy and Big Data: The Need for a Multi-Stakeholder Approach for Developing an Appropriate Privacy Regulation in the Age of Big Data", Ingo Pies, Halle, July 2015 [Available at SSRN: https://ssrn.com/abstract=2634970].

[46] H. Takabi, J. B. D. Joshi, G.-J. Ahn, "Security and Privacy Challenges in Cloud Computing Environments", IEEE Security and Privacy Journal, vol. 8, issue: 6, pp. 24-31, November 2010.

[47] A. Sachdev, M. Bhansali, "Enhancing Cloud Computing Security using AES Algorithm", International Journal of Computer Applications Journal, vol. 67, issue: 9, pp. 19-23, April 2013.

[48] H. T. Dinh, C. Lee, D. Niyato, P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches", Wireless Communications and Mobile Computing Journal, vol. 13, issue: 18, pp. 1587-1611, December 2013.

[49] N. Fernando, S. W. Loke, W. Rahayu, "Mobile cloud computing: A survey", Elsevier, Future Generation Computer Systems, vol. 29, issue: 1, pp. 84-106, January 2013.

[50] M. Ali, S. U. Khan, A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges", Elsevier, Information Sciences Journal, vol. 305, pp. 357-383, June 2015.

[51] S. Bhavani, A. Hatwal, U. Mittal, "Study on Cloud Computing and Different Load Balancing Algorithms in Cloud Computing", International Journal of Emerging Research in Management &Technology, vol. 4, issue: 5, pp. 331-336, May 2015.

[52] R. Iqbal, B. More, S. Mahmud, U. Yousuf, "Big Data analytics: Computational intelligence techniques and application areas", International Journal of Information Management, (in press), June 2016.

[53] X. Chang, Z. Ma, M. Lin, Y. Yang, A. G. Hauptmann, “Feature Interaction Augmented Sparse Learning for Fast Kinect Motion Detection”, IEEE Transactions on Image Processing, vol. 26, no. 8, pp. 3911-3920, August 2017.

[54] ITU, "H.265: High efficiency video coding," ITU, 9/7/2015. [Online]. Available: http://www.itu.int/rec/T-REC-H.265. [Accessed 12/10/2016].

[55] V. A. Memos, K. E. Psannis, “Encryption algorithm for efficient transmission of HEVC media”, Journal of Real-Time Image Processing, May 2015.

[56] A. Cotton, "Defining the Future of Television," BBC, 2/11/2015. [Online]. Available: http://www.bbc.co.uk/rd/blog/2013/06/defining-the-future-of-television. [Accessed 13/10/2016].

[57] D. C., J. Joseph, “Leading Television Industry Players Line Up To Support '4K Ultra HD”, Consumer Technology Association, 11/11/2014. [Online]. Available: https://www.cta.tech/News/Press-Releases/2014/November/Leading-Television-Industry-Players-Line-Up-To-Sup.aspx. [Accessed 12/10/2016].

[58] J. Lowensohn, “YouTube now supports 4k-resolution videos”,cnet, 9/7/2010. [Online]. Available: https://www.cnet.com/news/youtube-now-supports-4k-resolution-videos/. [Accessed 14/10/2016].

[59] X. Chang, Z. Ma, Y. Yang, Z. Zeng, A. G. Hauptmann, “Bi-Level Semantic Representation Analysis for Multimedia Event Detection”, IEEE Transactions on Cybernetics, vol. 47, no. 5, pp. 1180-1197, May 2017.

[60] S. A., G. Petrin, "Ultra High Definition Television: Threshold of a new age", ITU Committed to connecting the world, 24/5/2012. [Online]. Available: http://www.itu.int/net/pressoffice/press_releases/2012/31.aspx#.WAOc1clfae8. [Accessed 14/10/2016].

[61] M. Chacksfield, "UHDTV to be name for both 4K and 8K television standard?", techradar - The Source for Tech Buying Advice, 28/5/2012. [Online]. Available: http://www.techradar.com/news/television/uhdtv-to-be-name-for-both-4k-and-8k-television-standard-1082328. [Accessed 13/10/2016].

[62] LG, "LG Launches World’S First 84-Inch UD 3D TV with Unparallel Picture Quality", LG Newsroom, 22/8/2012. [Online]. Available: http://www.lgnewsroom.com/2012/08/lg-launches-worlds-first-84-inch-ud-3d-tv-with-unparallel-picture-quality/. [Accessed 14/10/2016].

[63] D. Quick, "Sony releases world's first 4K home theater projector", New Atlas - Sony, 31 5 2012. [Online]. Available: http://newatlas.com/sony-vpl-vw1000es-4k-projector/22760/. [Accessed 14/10/2016].

[64] Sony, "Immerse yourself in Sony 4K - Sony 4K home theater projector", Sony, 19/12/2011. [Online]. Available: https://dealersource.sel.sony.com/dsweb/p/builtin/sony_4k_home_theater.html. [Accessed 14/10/2016].

[65] I. A. T. Hashem,I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, S. U. Khan, "The rise of “Big Data” on cloud computing: Review and open research issues", Information Systems, vol. 47, pp. 98-115, January 2015.

[66] S. Zhang, H. Wang, W. Huang, Z. You, “Plant Diseased Leaf Segmentation and Recognition by Fusion of Superpixel, K-means and PHOG”, Optik-International Journal for Light and Electron Optics, vol. 157, November 2017.

[67] S. M. Abdulhamid, M. S. A. Latiff, S. H. H. Madni, M. Abdullahi, “Fault tolerance aware scheduling technique for cloud computing environment using dynamic clustering algorithm”, Neural Computing and Applications, pp. 1-15, July 2016.

[68] C. L. P. Chen, C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data", Elsevier, Information Sciences, vol. 275, p. 314–347, August 2014.

[69] B. B. Gupta, O. P. Badve, “Taxonomy of DoS and DDoS attacks and desirable defense mechanism in a cloud computing environment”, Neural Computing and Applications, vol. 28, issue 12, pp. 3655-3682, December 2017.

[70] X. Chang, Y.-L. Yu, Y. Yang, E. P. Xing, “Semantic Pooling for Complex Event Analysis in Untrimmed Videos”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 8, pp. 1617-1632, August 2017.

[71] I. Ahmad, M. A. Z. Raja, M. Bilal, F. Ashraf, “Neural network methods to solve the Lane–Emden type equations arising in thermodynamic studies of the spherical gas cloud model”, Neural Computing and Applications, vol. 28, supplement 1, pp.929-944, December 2017.

[72] A. Balasubramanian, R. Mahajan, A. Venkataramani, “Augmenting Mobile 3G Using WiFi: Measurement, System Design, and Implementation”, in ACM MobiSys ‘10, San Francisco, California, USA, 15-18 June 2010.

Chapter 13

Green Cloud Communication System for Big Data Management

C. L. Stergiou, K. E. Psannis, Y. Ishibashi, "Green Cloud Communication System for Big Data Management", in Proceedings of The 3rd World Symposium on Communication Engineering (WSCE 2020), 9-11 October 2020, held Online, Thessaloniki, Greece. [DOI 10.1109/WSCE51339.2020.9275579]

13.1 Introduction

This paper makes an effort to survey and study the open challenges in the field of energy-efficient and green Cloud infrastructures. The useful software that offers the possibility to implement and evaluate Cloud environments is CloudSim, which I also use to demonstrate and propose my idea. Moreover, I consider CloudSim's simulator architecture to achieve the proposal. Consequently, I investigate and propose a systematic framework for better use of Big Data management, based on a Cloud federated network. Additionally, I propose an algorithm for achieving an energy-efficient resource allocation technique for Big Data management in the Green Cloud environment. The experimental results demonstrate that our proposed model has immense potential as it offers significant performance gains regarding cost-saving and better data management under large workload scenarios.

13.2 Related Work

For the purpose of our work we study and analyze previous literature which has been contributed with Cloud simulators, and more specifically with CloudSim.

There are a few studies that are more relevant to our work. Initially, Louis et al. [4] propose CloudSimDisk, which is a scalable module for modeling and simulation of energy-aware storage in a cloud system. Authors work's contribution is in the field of Cloud Computing, aiming to extend the widely used CloudSim simulator. Also, Dr. R. Malhotra & P. Jain [5] initially define the use of CloudSim. Thereafter, they explore all the variants that are available in Cloud Simulators, such as CloudAnalyst, GreenCloud, Network CloudSim, EMUSIM and MDCSim. Consequently, Dr. R. Malhotra & P. Jain compare the use of all CloudSim Variant with respect to networking, platform and language. Moreover, Buyya et al. [6] propose three aspects: (a) Architectural principles for energy-efficient management of Clouds, (b) Energy-efficient resource allocation policies and scheduling algorithms with consider to the quality-of-service expectations and devices power usage characteristics, and (c) A novel software technology for energy-efficient management of Clouds. Furthermore, A. V. Sajitha & Dr. A. C. Subhajini [3] present a comprehensive information about default energy conscious Virtual Machine placement algorithms in the CloudSim. Through their experiments, they can conclude that CloudSim simulator is better than the others regarding the energy oriented Data Center evaluation, especially in dynamic environments. They reach this conclusion because CloudSim is an extensible open source software. Finally, Long et al. [7] introduce CloudSim simulator as a framework which provides simulation scenarios for power to manage services and modeling of cloud infrastructure. In addition to this, they analyze CloudSim's architecture, and how it could be used it in order to model a Cloud environment.

Papers **Features to explore**	**Louis et al. [4]**	**Dr. R. Malhotra & P. Jain [5]**	**Buyya et al. [6]**	**A. V. Sajitha & Dr. A. C. Subhajini [3]**	**Long et al. [7]**
Fr	**X**				**X**
Pl		**X**	**X**	**X**	
Ar			**X**	**X**	**X**
BDP			**X**	**X**	
BDA		**X**			**X**
BDM	**X**	**X**	**X**	**X**	**X**
IS		**X**	**X**		**X**
CSM	CloudSimDisk	Multiple Simulators	CloudSim: Energy-Aware Scenario	CloudSim: Energy Efficient Green Cloud Data Center	CloudSim: Energy Efficient Cloud Environment

Table 13.1: Related Work Comparison

Fr: Framework	**BDA:** Big Data Analytics
Pl: Platform	**BDM:** Big Data Management
Ar: Architecture	**IS:** Integration Scenario
DBP: Big Data Processing	**CSM:** CloudSim Model

Table 13.2: Features to Explore Abbreviations

Particularly, Table 13.1 illustrates the major features contributed more to our working scenario, as mentioned in previous works. Table 13.2 presents the meanings of the abbreviations listed in Table 13.1. Specifically, we can observe that each work demonstrates a different CloudSim scenario regarding the simulation expectations. Additionally, three of the related works perform a platform scenario, three perform an architecture scenario and only two of them perform a framework scenario. In addition to this, two of the related works perform a combine scenario of platform and architecture and one of them perform a combine scenario of framework and architecture. Moreover, all of them contribute Big Data Management scenario, in contrast of Big Data Analytics and Big Data Processing that contributed by only two of them. Finally, three of the related works illustrates an integrated technologies scenario.

13.3 Theoretical Approach

13.3.1. CloudSim Simulator Architecture

CloudSim simulator provides a plenty of layers that works on them. Its layers support the modeling and the simulation of a Cloud environment. It also offers the ability of setting requirements about memory, storage and bandwidth parameters of both VMs and Cloud servers. The whole layer setup of CloudSim is shown in figure 13.1. Due to figure 13.1 and the architecture supported from CloudSim, a Cloud

Service Provider (CSP) has the ability to evaluate a customized method based on these layers in order to verify and competitive the various policies of VM provisioning [8].

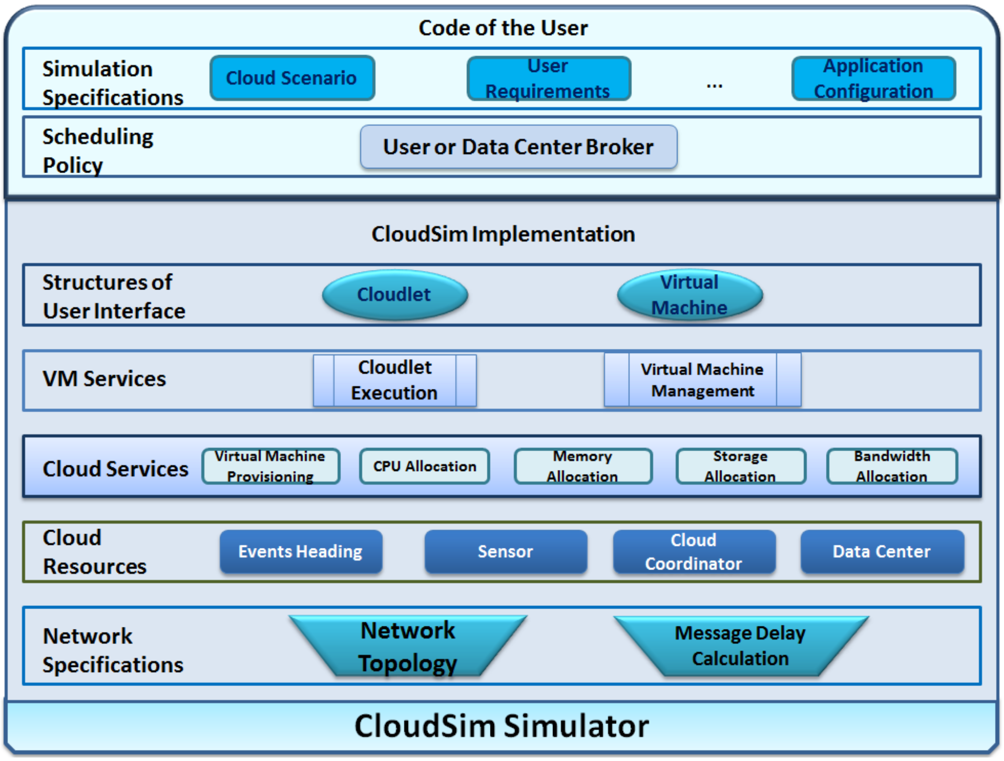


Figure 13.1: CloudSim Architecture with layers.

13.3.2. Data center energy efficiency - Green Grid Consortium

Finally, there is another scenario, due to “Green Grid Consortium”, in order to calculate the Power Usage Effectiveness (PUE). This could be delivered by dividing the IT Equipment Power (ITEP) which demonstrates the energy facilities consumed from the equipment which is used in order to manage, process, transfer, store, operate, and route data through the data, with Total Facility Power (TFP) which demonstrates the data centers’ entirely power that is delivered. This is an inverse version of PUE calculation, described as Data Center’s Infrastructure Efficiency (DCIE), and illustrated by Equation (1).

$$DCIE = \frac{ITEP}{TFP} \qquad (1)$$

Count on the complexity of each Cloud environment and the forceful nature each resource demanding, Reinforcement Learning (RL) scenario could be utilized in order to suggest optimal allocation rules.

13.3.3. Modeling Power Consumption

Several works which have been made in this field [9] [10] show that applying Dynamic Voltage and Frequency Scaling (DVFS) on a CPU could result in nearly linear power to frequency relationship. This fact based both on the restricted number of the states that could place the value of frequency and voltage of a CPU, and also on the reason that the DVFS applied only on CPU and not to any other hardware component. Furthermore, related works exhibit that the consumption of an idle server is approximately in average of 75% power consumption of a server operating at maximum CPU speed. Regarding this, the switching of idle servers off aiming to reduce the overall Power Consumption is justified. This justification concludes in the definition of Equation (2) bellow:

$$P_s = pf * PC_m + (1 - pf) * PC_m * CPU_u \qquad (2)$$

In Equation (2), PC_m represents the maximum value of Power Consumption, in the moment where the server is fully utilized. Then, pf illustrates the value of the Power Consumed Fraction of the idle server, and CPU_u represents the utilization of the CPU. CPU's utilization might be changed during the time duration count on the variability of each process workload.

Consequently, the utilization of the CPU could also be demonstrated as a function of time. This can be represented as $CPU_u(t)$. Hence, the overall energy (OE) consumed by a settled node could be defined from Equation (3) bellow, as an integral part of the power consumption function for a period of time:

$$OE = \int_{time} P_s(CPU_u(t)) \qquad (3)$$

13.3.4. Cloud Federation approach

CSPs run various data centers, in many places simultaneously, through the Internet aiming to achieve the needs of multiple customers all over the world. Therefore, their systems which are currently established do not have the ability to support policies and mechanisms for dynamically coordinating load-shredding between various data centers with the aim to outline an optimal location in order to host application services and achieve rational QoS levels.

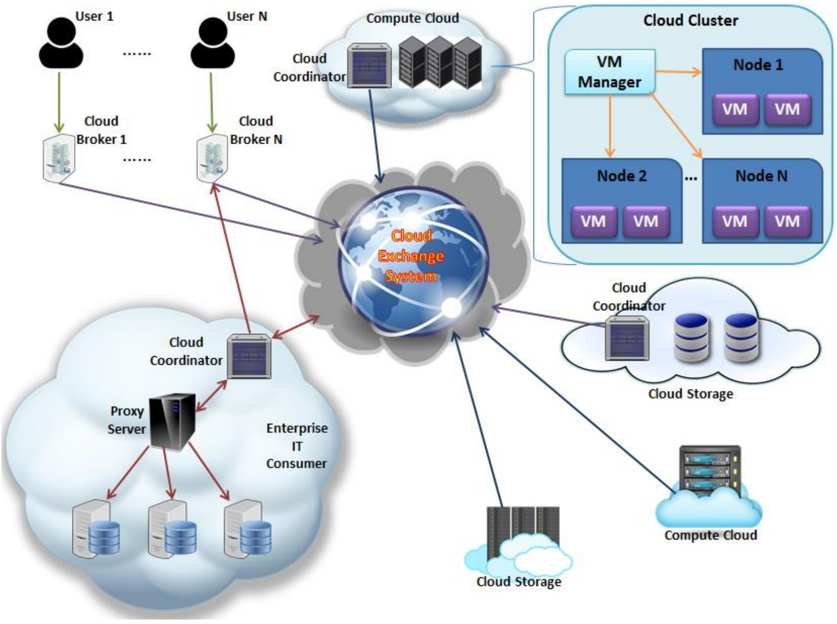


Figure 13.2: Federated Cloud Network.

Figure 13.2 illustrates a Cloud Computing system architecture consisting of services for different users' needs. These needs count on Cloud SaaS model, brokering, and CSPs coordinator services, which could be able to boost utility-driven Cloud services, such as application provisioning and workload migration.

A federated Internet-based system of supervisory distributed Cloud, which could provide particular benefits both in finance and performance. Such benefits could be the improvement of SaaS model's ability by achieving the QoS levels, and also the enhance of the peak-load of data that could be handled and the dynamic expansion of storage capacity to every user that could have access to the Cloud system. Thus, the Cloud federated system could enlarge the reliability of the CSPs participation except of ensuring the business continuity.

Cloud Coordinator (CCo) is one of the major components of Federated Cloud architecture. CCo is established in every Cloud system and it has the main responsibilities of system's operation. CCo could export Cloud services in whole federated infrastructure, and also has the ability to track the data load through the Cloud resources and then undertaking negotiation with the other CSPs established in the federation aiming to handle the sudden peak in the resource demand at each local Cloud system. In addition to this, CCo could monitor every application's execution and their lifecycle.

13.4 Experimental Setup & Evaluation

In this section we try to clarify the way of thinking that led to the parameters set for the proposed scenario.

13.4.1. Data Center Configuration Setup

To our scenario, the data center needs to be created with at most 250 heterogeneous host machines, and at least 50 heterogeneous host machines. For each of them two types of server configurations are used as shown in Table 13.3.

Host Type 1	Host Type 2
HP ProLiant ML110 G4	HP ProLiant ML110 G5
Intel Xeon 3040	Intel Xeon 3075
Cores: 2 (1800MHz/core)	Cores: 2 (2200MHz/core)
4GB RAM memory	4GB RAM memory
1TB storage memory	1TB storage memory

Table 13.3: Server Configurations

13.4.2. Experimental Configuration Setup

The whole simulation analysis contacted on the two aforementioned types of hosts, randomly selected, for respectively 50, 100, 150, 200, and 250 heterogeneous VMs each time. For each VM the configuration setup illustrated in Table 13.4.

Virtual Machine
1 x CPU (with 1 core)
1000 or 2000 MIPS capacity (randomly)
8GB RAM memory
2TB storage memory

Table 13.4: VMs Configurations

The value of power consumed by the VMs is defined according our proposed model. Model's host consumption is set from 150Watt on 0% CPU utilization up to 300Watt on 100% CPU utilization. Each user who contacted the system is considered as a different independent user, who submit its requests to the host for provisioning of 50, 100, 150, 200, or 250 VMs that fills the full capacity of data center. The VM simulated to our scenario runs a web application with variable workload value, which affects to CPU's utilization. The web application runs to each VM operates for 150000 MIPS. The whole experiment runs for 7 days.

13.5 Experimental Results

In this section, the experimental results of the simulation scenarios implemented on CloudSim simulator along with the other findings are demonstrated and analyzed.

13.5.1. Algorithm Approach

Through this work we propose an algorithm scenario that embeds novel techniques for energy-efficient Big Data management in CloudSim toolkit.

Reinforcement and Federated approaches are combined in our model in order to save computation power and distribute it to multiple CSPs.

As we already know from the literature, CloudSim could be used as a virtualization software aiming to simulate and operate multiple scenarios for the function of Cloud Data Center [10]. Also, CloudSim could be used as a framework for scalable simulation process which facilitates support scenarios and experimentations of virtual data centers in Cloud environments. As regards the data management, it offers scenarios of services for VMs, based on memory, storage and bandwidth, with multiple configurations. These scenarios could be established without limitation on data volume, and as a result could be operated on large-scale data.

Algorithm 1

```
CloudSim impute values:
    ➢ nH: Total number of the various Hosts of data centers.
    ➢ nVM: Total number of the Virtual Machines used in the operation process.
    ➢ WC: The value of Workload Consumption of CPU's power.
    ➢ PF: The value of Power Consumed Fraction of idle server.
    ➢ cu: The value of CPU's utilization.
    ➢ N: The maximum value of process operation.
CloudSim outcomes:
    ✓ OA: Overall Allocation of the VMs use
    ✓ OE: Overall Energy consumed by entire process
Proposed Scenario:
initialize nH, nVM, WC // create multiple Hosts, VMs and Workload  Consumption of
                          CPU's power
initialize Cloud Environment() // initialize state: S, effect: E, and Y values and X
                                  counter values
for (VM ϵ nVM) and (Host ϵ nH)
    S = {WC, h, vm}
    for i=0, i<=N, i++
        $E_i = E \epsilon \max_E * Y_i(S_i, E)$
        with $E_i$ count $E_{i+1}$
        recompense $X_{i+1}$
        $Y_{i+1}(S_i, E_i) \leftarrow X + [WC \cdot \max_E \cdot Y_i(S_{i+1}, E_{i-1})]$
        $S_i = S_{i+1}$
    end
    $OE_N = (PF * WC + (S_i * Y_{i-1})) * cu_N$
    return h
    distribute(Host, VM)
end
return OA and OE
```

Regarding the proposed scenario, through CloudSim toolkit we can perform a better use and performance of the associate obstacles in order to achieve better services and policies based on data management techniques. Count on the existing CloudSim architecture, we could support Cloud data centers infrastructures with the Data Center's Infrastructure Efficiency (DCIE) applied on it. Therefore, our proposed

CloudSim scenario illustrates a framework which encompasses all the factors needed and integrated in order to achieve an efficient and Green Cloud environment, based on the existing CloudSim architecture.

13.5.2. DCIE resource allocation results

The DCIE emphasis metric and its affect are produced over the 50% of overall scenario level. The proposed scenario contacted for seven days duration, for four dimension of the number of each VM.

Figure 13.3 demonstrates the DCIE metric comparison of our proposed scenario. Due to figure 13.3, the DCIE value of resource allocation shows that the Day 1 has the highest value of DCIE, in contrast with Day 7 which has the lowest value of DCIE. Moreover, the lowest value for each category of VMs are: a) for 50 VMs is approximately 47%, b) for 100 VMs is approximately 50%, c) for 150 VMs is approximately 52%, d) for 200 VMs is approximately 54%, and d) for 250 VMs is approximately 62%. Count on the aforementioned results, the value of DCIE increases by approximately 2% per day. As regards the grow of the value of DCIE, the impact of this value to the infrastructure efficiency of the data center, summarized from the proposed algorithm, additionally guarantees and provides such a good and efficient infrastructure environment.

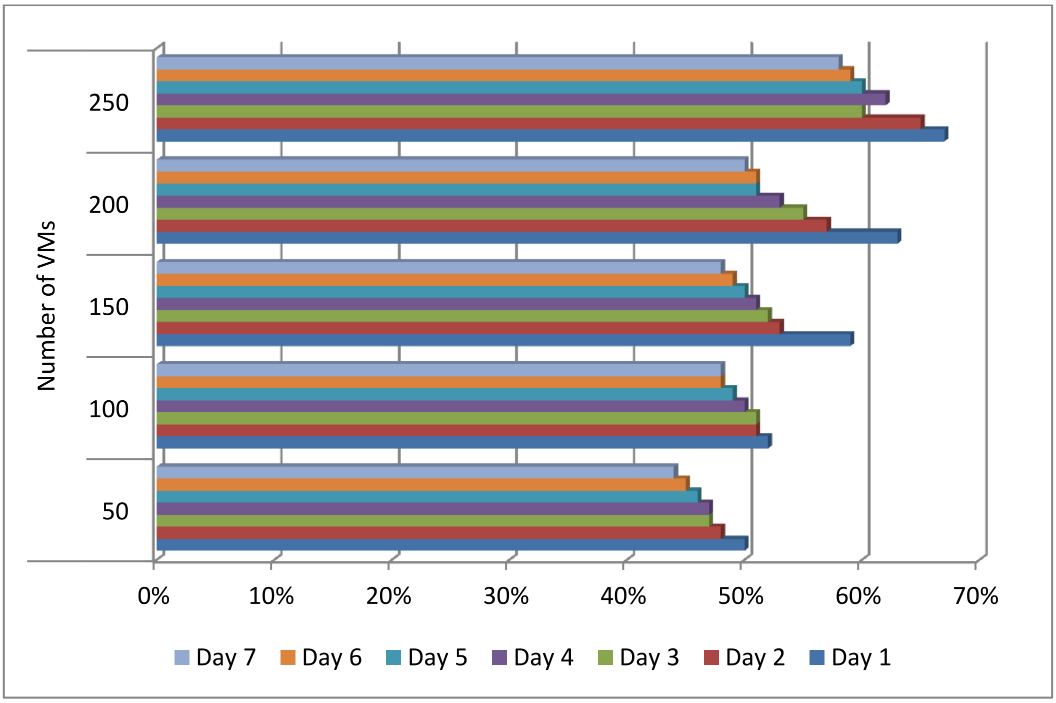


Figure 13.3: Metric of DCIE count on proposed scenario

13.5.3. Performance analysis

The whole power consumed of the system could be delivered by multiple figures analyzed in the current subsection. The factors that we take into account are Power

Consumed Fraction, CPU's utilization, Overall Energy consumed, and Data Center's Infrastructure Efficiency.

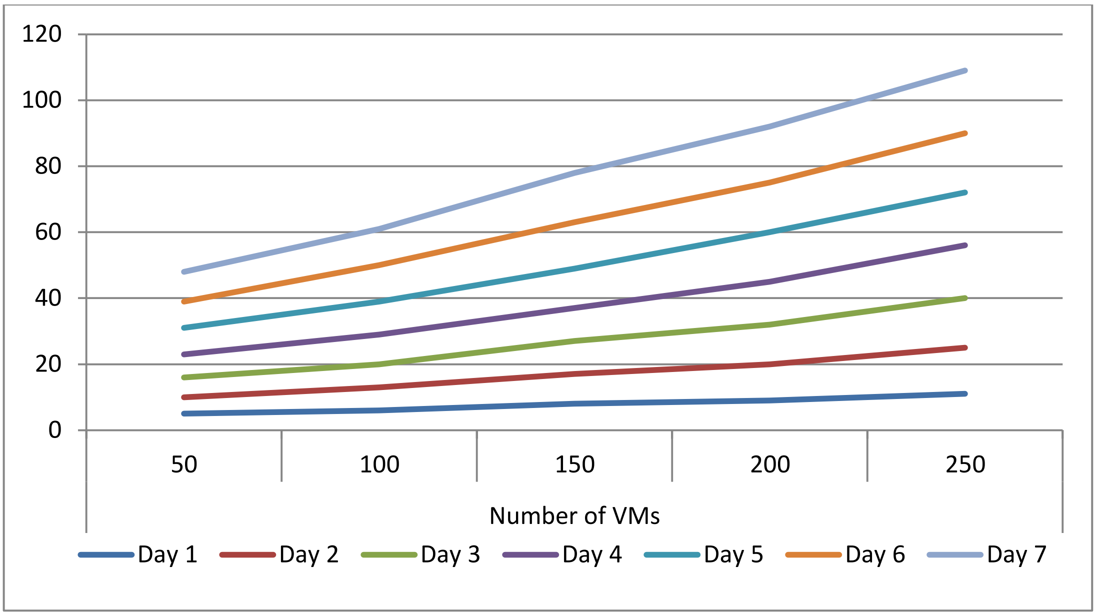


Figure 13.4: Power Consumed Fraction by energy-efficient algorithm

Figure 13.4 presents the effects of Power Consumed Fraction of the simulation regarding the different need of VMs, over the seven days of the procedure. In particular, we can define that as the days past the need of power consumed rises. Additionally, as many VMs we contribute such many power resources needed.

13.6 Chapter Summary

After we ended up using the CloudSim simulator as the ideal Cloud simulator for our scenario, we investigated the related works on the field that try to offer energy-efficient and green Cloud environment for data management and processing through the CloudSim simulator. CloudSim could offer aspects such VMs and CCo that established and used equally in real-time scenarios and infrastructures. Based on our study, we proposed a system framework for better use of Big Data management, based on Cloud federated network. Furthermore, we tried to offer an algorithm approach of our scenario, by proposing a novel model for achieving an energy-efficient resource allocation technique for Big Data management on Green Cloud environment. Finally, we demonstrated experimental results which show that our proposed model has immense potential as it offers significant performance gains regarding the cost saving and the better data management under large workload scenarios.

For the future we are oriented to involve to the proposed framework the major issue of security, with the aim to provide a Green and Secure Cloud environment, through the scenarios of federated learning and reinforcement learning.

13.7 Chapter References

[3] A. V. Sajitha, Dr. A. C. Subhajini, “Analysis of Cloud Sim Toolkit for Implementing Energy Efficient Green Cloud Data Centers”, International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 6, issue IV, pp. 4613-4624, April 2018.

[4] B. Louis, K. Mitra, S. Saguna, C. Ahlund, “CloudSimDisk: Energy-Aware Storage Simulation in CloudSim”, in Proceedings of 2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC), Limassol, Cyprus, 7-10 December 2015, pp. 11-15. [DOI: 10.1109/UCC.2015.15]

[5] Dr. R. Malhotra, P. Jain, “Study and Comparison of CloudSim Simulators in the Cloud Computing”, The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), vol. 1, no. 4, pp. 1-5, October 2013. [DOI: 10.9756/sijcsea/v1i4/0104510201]

[6] R. Buyya, A. Beloglazov, J. Abawajy, “Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges”, in Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2010), 12-15 July 2010, Las Vegas, USA. [arXiv:1006.0308]

[7] W. Long, L. Yuqing, X. Qingxin, “Using CloudSim to Model and Simulate Cloud Computing Environment”, in Proceedings of 2013 Ninth International Conference on Computational Intelligence and Security, 14-15 December 2013, Leshan, China.

[8] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, R. Buyya, “CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms”, Wiley, Journal of Softwaer: Practice and Experience, vol. 41, issue: 1, pp. 23-50, January 2011. [DOI: 10.1002/spe.995]

[9] A. Gandhi, M. Harchol-Batler, R. Das, C. Lefurgy, “Optimal power allocation in server farms”, in Proceedings of ACM 11th international joint conference on Measurement and modeling of computer systems, pp. 157-168, June 2009, Seattle, WA, USA.

[10] R. Raghavendra. P. Ranganathan, V. Talwar, Z. Wang, X. Zhu, “No "power" struggles: coordinated multi-level power management for the data center”, ACM SIGARCH Computer Architecture News, vol. 36, no. 1, March 2008. [DOI: 10.1145/1353534.1346289]

Chapter 14

InFeMo: Flexible Big Data management through a federated Cloud system

C. L. Stergiou, K. E. Psannis, B. B. Gupta, "InFeMo: Flexible Big Data management through a federated Cloud system", ACM Transactions on Internet Technology, In Press, 2020.

14.1 Introduction

This paper introduces and describes a novel architecture scenario based on Cloud Computing and count on the innovative model of Federated Learning. The proposed model is named *Integrated Federated Model*, with the acronym *InFeMo*. InFeMo incorporates all the existing Cloud models with a federated learning scenario, as well as other related technologies that may have integrated use with each other, offering a novel integrated scenario. In addition to this, the proposed model is motivated to deliver a more energy-efficient system architecture and environment for the users, which aims to the scope of data management. Also, by applying the InFeMo the user would have less waiting time in every procedure queue. The proposed system was built on the resources made available by Cloud Service Providers (CSPs), by using the PaaS (Platform as a Service) model, to be able to handle user requests better and faster. This research tries to fill a scientific gap in the field of federated Cloud systems. Thus, taking advantage of the existing scenarios of FedAvg and CO-OP, I keen to end up with a new federated scenario that merges these two algorithms, and aiming to has a more efficient model, that it can select, depending on the occasion, if it “train” the model locally in the client or globally in server.

14.2 Related Review

14.2.1. Big Data Management in Cloud

During the last years, several works have been made in order to manage data, and more specificly Big Data, in Cloud environments. Thus, for the purpose of this research we have studied and analyzed previous literature researches that have been made in the field of data management in Cloud environment [18-23]. The following paragraphs present the previous to our study research papers.

To begin with, Thakur et al. [18] present a Robust reputation management mechanism, that tries to encourage the Cloud Providers (CPs) in a federated cloud to separate users between good and malicious, and grant resources in such a way that they do not share them.

Moreover, Cai et al. [19] present a novel in-memory data management system, called Memepic, that unifies both online data query and data analytics functionality, permitting low-latency storage service and efficient in-situ data analytics.

Another work in this field is presented by Pasquier et al. [20], which introduce Information Flow Control (IFC) model and describe and evaluate this IFC architecture and implementation (CamFlow) that compromises an OS level execution of IFC with support for application management, in cooperation with an IFC-enabled middleware.

To continue with, Zhu et al. [21] present a controllable blockchain data management (CBDM) model that can be deployed in a Cloud environment, which it can evaluate its security and performance, in order to demonstrate utility.

A heterogeneous data storage management scheme that flexibly provides simultaneously deduplication management and access control over innumerable CSPs is introduced by Yan et al. [22].

Finally, a trust based federated identity management as a Cloud based utility service is presented by Premarathne et al. [23]. Furthermore, this work proposes a Cloud-based utility service model for federated identity-based trust negotiations management and a novel trust based evaluation method to access the cooperativeness of the identity providers to improve the reliability.

14.2.2. Federated Learning Scenarios

On the other hand, there are a number of remarkable works associated with the novel scenarios of Federated Learning Systems [14-16] [24-28]. The following paragraphs present the relative research papers.

Initially, a general distributed multiquery processing problem motivated by the need to speedup data acquisition in federated databases using evolutionary algorithm presented by Mansha & Kamiran [24].

Furthermore, Yao et al. [14] through the experiments presented in their work show that baseline methods could outperformed by their proposed model, especially in Non-IID data distributions, and accomplishes a compression of more than 20% in required communication rounds.

In another related work, Wang et al. [25] propose an algorithm that adapts to real-time system dynamics, derived from theoretical analysis. Then, it defines that every specific iteration contains a local update step which is possibly followed by a global aggregation step.

Nilsson et al. [15] benchmarks three federated learning algorithms *(1) Federated Averaging (FedAvg)* [26], *(2) Federated Stochastic Variance Reduced Gradient (FSVRG)* [26], *(3) CO-OP* [27] and compare their performance opposed to a centralized technique in which data resides on the server.

Moreover, Young et al. [28] presents an approach to computing the covariance matrix with federated databases. Also, it computes the exact covariance matrix rather than an approximation.

Finally, McMahan et al. [16] advocate a different scenario that leaves the training data distributed on the mobile devices, and learns a shared model by aggregating locally-computed updates. Thus, this work introduces the Federated

Averaging algorithm, which integrates local stochastic gradient descent (SGD) on each client with a server that performs model averaging.

14.3 System Formulation & Analysis

The first goal of this work is to introduce a flexible data management system that it is set up in an academic server and it operates with the useful help of the cooperative Cloud Providers. The system set up does not differ to any other structure of server set regarding the hardware. The main purpose is to be set a cooperative system, a federated scenario, in which the cooperative CSPs could "help" the load-balance of data management and transmission by having the load of user authentication.

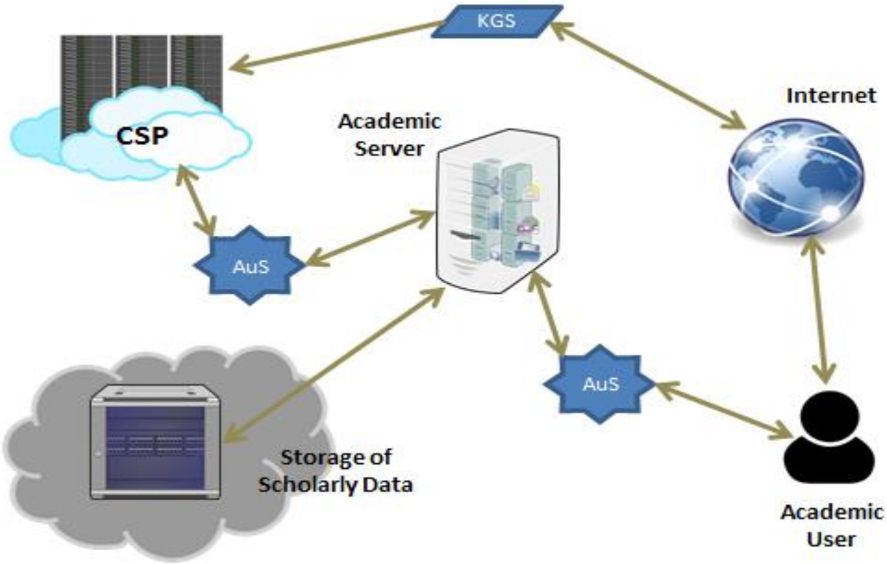


Figure 14.1: System architecture.

Figure 14.1 shows the operation of our proposed scenario implementation. More specifically, the academic user will be able to access academic data via a CSP (Cloud Service Provider) that is authorized to access the server of the academic institution. This will enable instantly and faster the data management through the benefits of the CSP Cloud platform.

The user authentication will be done in two parts. Initially, through the CSP with the KG (Key Generation System) that will give the unique key until the connection expires. Then, it will be authenticated through the academic server, where the user will be authenticated, as well as the level of permissions granted to the user.

There will be some states of rights: 1) Data owner, 2) Data researcher, 3) student. Depending on the property that results from the authentication process taking place on the CSP platform, the level of data management will be obtained.

14.3.1. Evaluation Approach Scenario

Depending on the system introduced above, we can clearly explain its function with figure 14.2. More specifically, we will try to explain the operation of the proposed system when multiple users try to have access to the Academic Server.

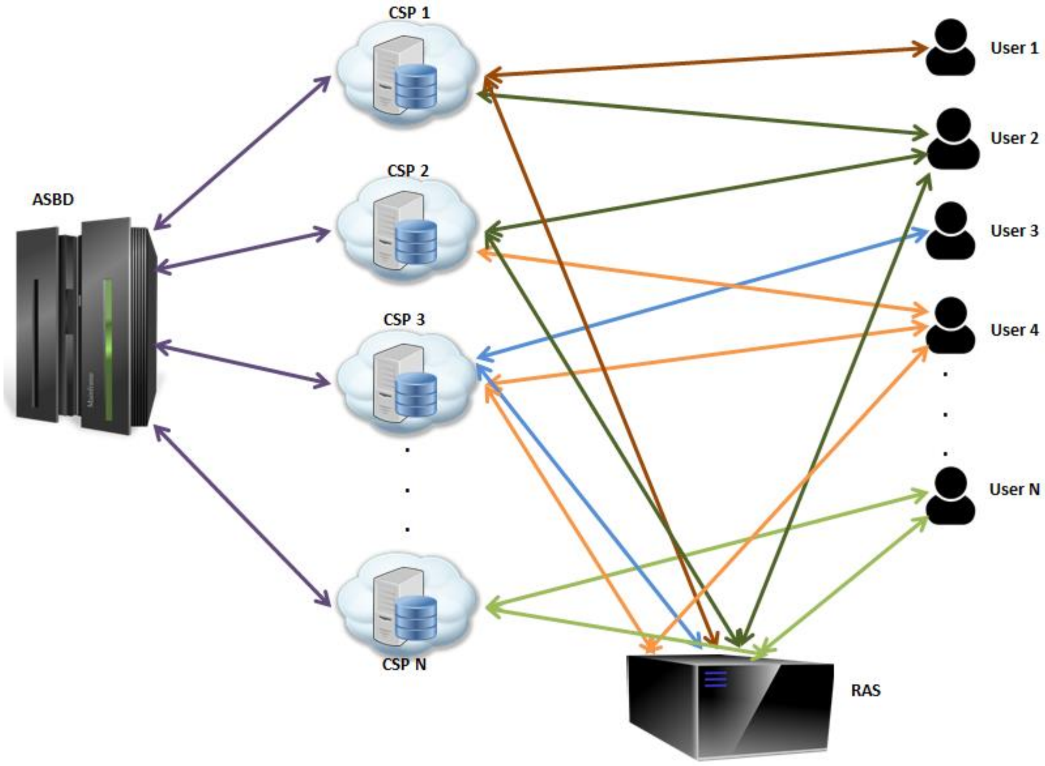


Figure 14.2: Academic Server evaluation and procedure.

Figure 14.2 show the operation and procedure of the Academic Server evaluation scenario. Accurately, the server follows the exact steps in its operation:

1) Connection request to CSP (1 or 2 or 3).
2) Control - User authentication in RAS (Reputation Authentication System)
3) Response of RAS through the open communication channel, directly to the user.
4) Depending on the user permissions level (1 or 2 or 3) the corresponding CSP data center will be used.
5) Communication of the CSP with the ASBD (Academic - Scholarly Big Data) repository to open a direct user communication channel.

The combination of 1 to 5 depends on the demands of each user and in the available Cloud provider at the moment of the user's request. Then, depending on the internet connection and the availability of the CP, the RAS authenticates the user and the level of user's permissions to the Academic Server. With this proposed system the user authentication takes place in the federated CSP environment and as a result the Academic Server is not burdened with this load. Thus, we can achieve an energy and computational efficient scenario for an academic server that will serve thousands of users, with different demands each one.

The whole proposed system set up is based on a federated learning system between the multiple CSPs, due to the general principle of federated learning systems. The general principle composed in training local models on local data samples and exchanging parameters between various local models at some frequency in order to generate a wide-spread model.

14.4 Scheme Implementation

In this section we try to introduce with more details the operation of the proposed system and the implementation design of it. All the information pertaining to the system's functionality are revealed in the following subsections.

14.4.1. Fundamental Procedure Scenarios

In this subsection, we introduce a number of fundamental algorithms of the proposed system. The operation of the Fundamental Procedure Scenario took part on an academic server for academic users. The whole system illustrates how some major processes could be completed based on our proposed scenario.

14.4.1.1. Data management through access by different CSPs.

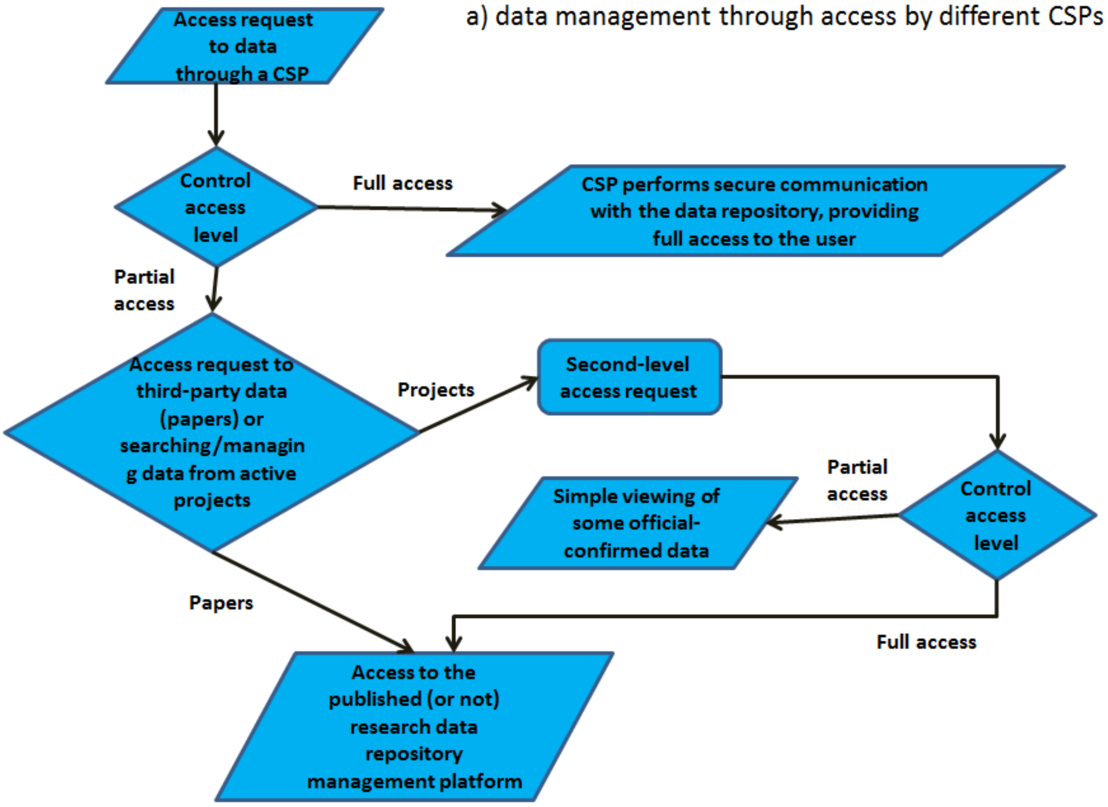

a) data management through access by different CSPs
Access request to data through a CSP
Control access level
Full access
CSP performs secure communication with the data repository, providing full access to the user
Partial access
Access request to third-party data (papers) or searching/managin g data from active projects
Projects
Second-level access request
Partial access
Control access level
Simple viewing of some official-confirmed data
Papers
Full access
Access to the published (or not) research data repository management platform

Figure 14.3: Data management through access by different CSPs.

Figure 14.3 shows the procedure of the management of data through the access by using multiple CSPs. Each step demonstrated in figure 14.3 is analyzed more clearly in the following steps, which depict how an academic user could request access to data and how this user could manage these data depending on the rights granted to the respective user. Particularly, the procedure follows the following steps:

Step 1 - The user makes a request for data access through a CSP.

Step 2 - A check is made on the CSP for the level of user access. There are two cases here, one is the user to have full access to the data, and the other is the user to have partial access to the data.

Step 3 - Full access - The CSP performs a secure communication channel with the data repository (academic server), providing full access to the user.

Step 3 - Partial access - A new request is sent by the user through the CSP for access to third-party data (scientific papers and works), or data from search / management related to participation in active projects. There are two cases here, one is related to the access to scientific papers and works, and the other is related to the access to data from finished and open academic projects.

Step 4 - Scientific papers & works - Access to the repository research platform of published research data located on the academic server.

Step 4 - Finished & open academic projects - A new second-level access request is made to authenticate the user and their rights.

Step 5 - A new second-level access level control for user rights is performed. There are two cases here, the one concern the full access of the user to the data from finished and open academic projects in the academic server, and the other concerns the partial access of the user to the data from finished and open academic projects in the academic server.

Step 6 - Full access - Provided by the academic server, through the CSP, full user access, and user access to the management platform.

Step 6 - Partial access - It is provided by the academic server, through the CSP, a simple view of certain, official-confirmed data to the user, from the management platform.

14.4.1.2. Deletion of data through access by different CSPs

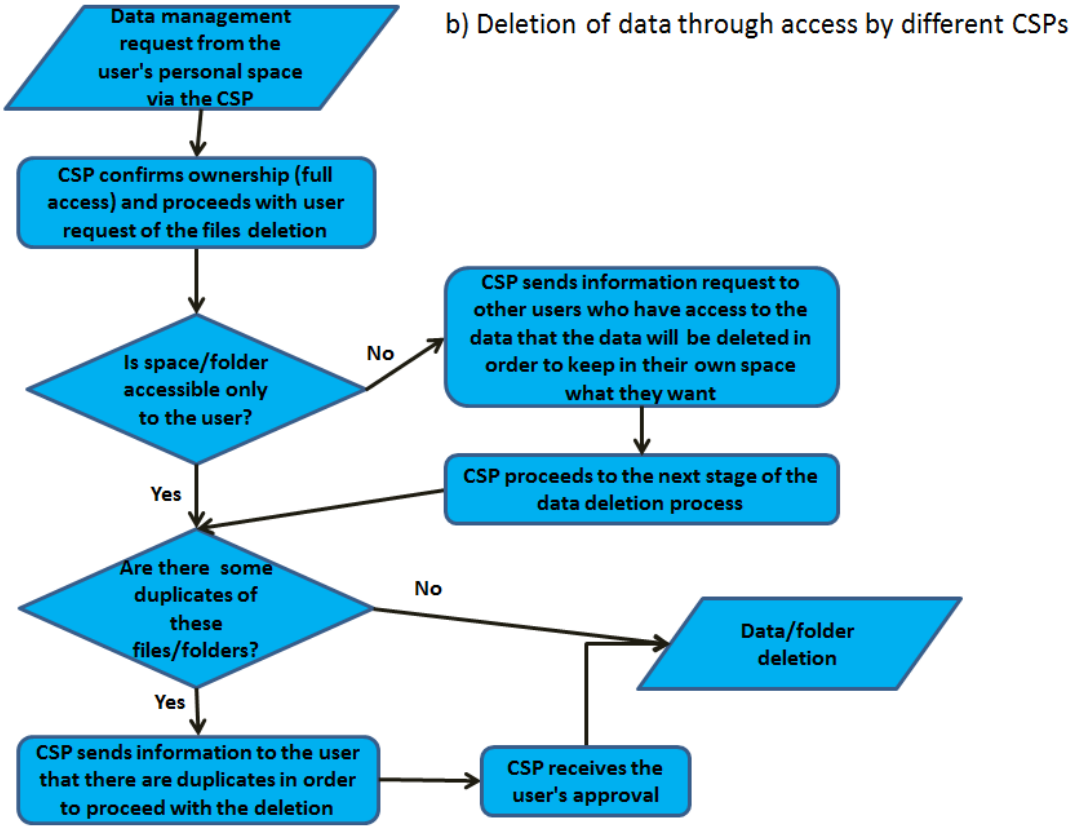

b) Deletion of data through access by different CSPs
Data management request from the user's personal space via the CSP
CSP confirms ownership (full access) and proceeds with user request of the files deletion
Is space/folder accessible only to the user?
No
CSP sends information request to other users who have access to the data that the data will be deleted in order to keep in their own space what they want
CSP proceeds to the next stage of the data deletion process
Yes
Are there some duplicates of these files/folders?
No
Data/folder deletion
Yes
CSP sends information to the user that there are duplicates in order to proceed with the deletion
CSP receives the user's approval

Figure 14.4: Deletion of data through access by different CSPs.

Figure 14.4 shows the procedure of deleting data on the academic server, through access from multiple CSPs. Each step demonstrated in figure 14.4 is analyzed more clearly in the following steps, which depict how an academic user could access data through the academic server and how this user could manage in order to delete these data depending on the rights granted to the respective user. Particularly, the procedure follows the following steps:

Step 1 - Request to delete data from the user in his/her personal space via the CSP.

Step 2 - The CSP confirms to the user that he has full access to and ownership of the content, and then deletes the data requested by the user.

Step 3 - User data rights control is performed. These are data/files only accessed by the user who requested the deletion or accessed by other users. There are two cases here, yes and no.

Step 4 - No - The CSP sends an update request to other users who have access to the data/files that the data/files will be deleted so that if they wish to keep copies in their own space on the system.

Step 5 - No - The CSP proceeds to the deletion of data/files. (*Then, the procedure follows the next step in a row: Step 4 - Yes*)

Step 4 - Yes - A data management system checks if there are any duplicates of the data/files that were requested to be deleted. Two cases arise in this step, yes and no.

Step 5 - No - Files are deleted from the file system.

Step 5 - Yes - The CSP sends the user information that there are duplicates elsewhere in the file system, and that it will delete them too.

Step 6 - After receiving approval from the user that it has been updated and accepts the duplicate deletion, it proceeds to delete the duplicate files from the system.

14.4.1.3. Adding data through access by different CSPs

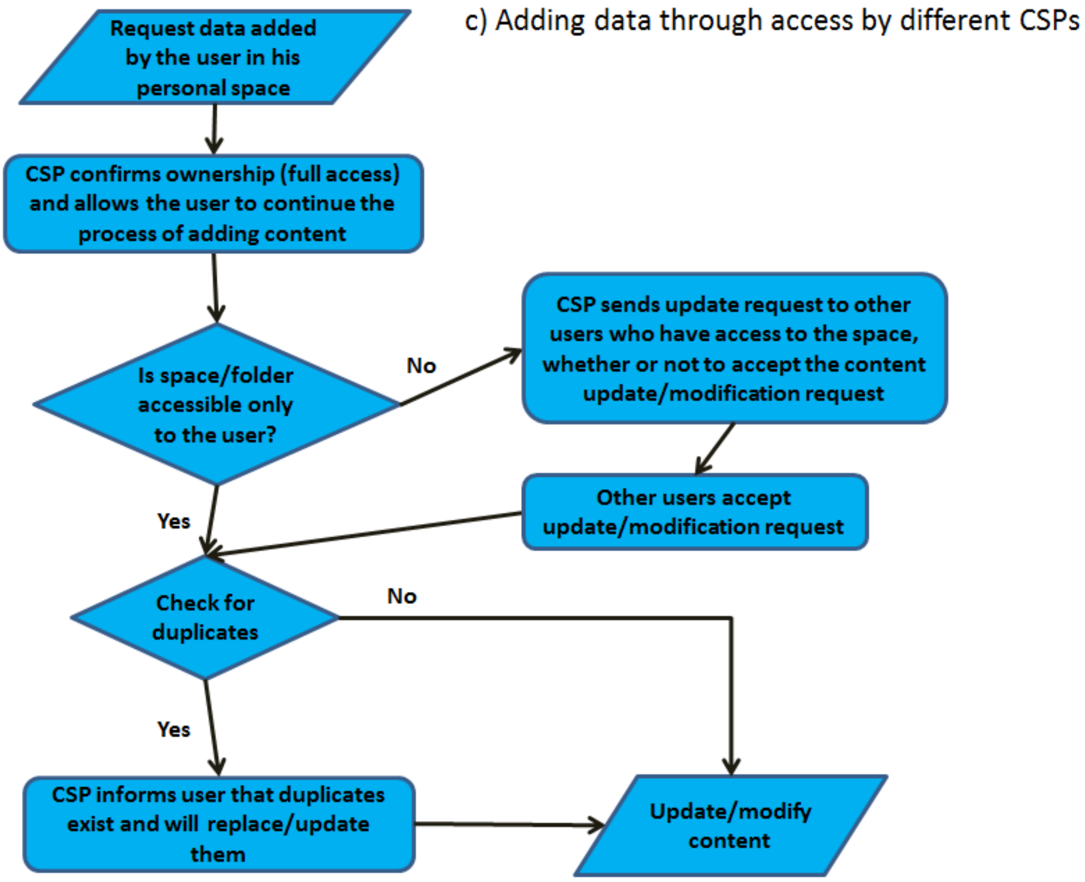

c) Adding data through access by different CSPs
Request data added by the user in his personal space
CSP confirms ownership (full access) and allows the user to continue the process of adding content
Is space/folder accessible only to the user?
No
CSP sends update request to other users who have access to the space, whether or not to accept the content update/modification request
Other users accept update/modification request
Yes
Check for duplicates
No
Yes
CSP informs user that duplicates exist and will replace/update them
Update/modify content

Figure 14.5: Adding data through access by different CSPs.

Figure 14.5 shows the procedure of adding data to the file system through access from multiple CSPs. Each step demonstrated in figure 14.5 is analyzed more clearly in the following steps, which depict how an academic user could access the owned disk space on the academic server and how this user could add an manage these data depending on the rights granted to the respective user. More particular, the procedure follows the following steps:

Step 1 - User request to add data to their personal space in the CSP system.

Step 2 - The CSP confirms authorized content ownership (full user access), and allows the content to be added by the authorized user.

Step 3 - The system checks whether the space the user is trying to modify/add content is a space/folder that is only accessible by the user or additional from someone else. Two cases arise in this step, yes and no.

Step 4 - No - The CSP sends an update (update request) to other users who have access to that space/folder for acceptance / approval / update access / content modification. (*Then, the procedure follows the next step in a row: Step 4 - Yes*)

Step 4 - Yes - Checking the file system, if any files already exist in the file system that the user wants to add (duplicate check). Two cases arise in this step, yes and no.

Step 5 - Yes - The CSP informs the user that duplicates exist and will replace/update existing ones with the new ones.

Step 6/Step 5 - No - Updating space/folder content.

14.4.2. Model Explanation

The operation of our proposed system is presented in this subsection. Generally, through the traffic that the central server receives from the requests of the various CSPs coming from the users, a system can be developed to support the communication of the CSPs with the academic server in order to follow the safest and fastest authentication methods through federated methods. The result of this process will be to allow the user to select the most appropriate CSP for the task the user wants to perform, and to develop a machine learning system through the communication of the CSPs with the academic server. This system will build on the resources made available by CSPs (using the PaaS - Platform as a Service) format to be able to handle user requests better and faster. This will develop a more efficient management system in a federated Cloud environment.

Thus, the user could have a more immediate communication with the academic server through the safe environment provided by the cooperative CSPs. Assuming the user is a client (k) that have contacted the server several times for a specific folder containing various type of data (n_k), so the cooperative CSPs could learn a scenario about this user in order to make the authentication procedure instantly and to navigate the user exactly to the most used files. This learning method could be established in the edge of communication of the each client and the collaborated CSPs. As a result, the academic server could reduce the computational ability for user services and focus on the most important, research process. Additionally, based on the 3 procedures presented in subsection 14.4.1, the system could also be applied to the training process taking place in the CSPs on federated learning concept. Thus, the users will be able to bypass some of the "easy" steps after a continuous and advanced use.

Moreover, the proposed Cloud model collaborates very well with the academic server and the various clients. The SaaS model assists in the storage of the amounts of data in the academic server through the assistance of the multiple collaborative CSPs. Also, the IaaS model is used due to the interface scenario that consists of the communication of the users with the academic server through the assistance of the multiple collaborative CSPs.

14.5 Algorithm Approach

The evaluation of our work can be additionally demonstrated through an algorithm analysis is inferred from the synchronous and asynchronous algorithms of federated learning scenarios. Studying the literature in the field of federated learning we ended up with two predominant algorithmic models, one of each category. The one is *Federated Averaging (FedAvg)* algorithm which is a synchronous algorithm and the other is *CO-OP* algorithm which is an asynchronous algorithm.

14.5.1. Federated Averaging (FedAvg) Algorithm

The Federated Averaging algorithm, or as briefly mentioned FedAvg, was initially introduced by A. Nilsson et al. [16]. This algorithm orchestrates training through a central server which hosts the shared global model w_t, where t is the communication round. Nevertheless, the actual optimization is done locally on clients using, for example, the Stochastic Gradient Decent (SGD). Moreover, FedAvg algorithm has five “*hyper-parameters*”, directly related to the general federated learning parameters, previously mention in the Introduction Section.

The parameters B, E, η, and λ are commonly used when training with SGD [17]. However, in FedAvg algorithm the variable E stands for the total number of iterations through the local data *before* the global model is updated [15].

In its operation, Federated Averaging algorithm begins with randomly initializing the global model of w_0. Specifically, one communication round of FedAvg algorithm drives to the consisting of the following aspect: (*algorithm operation procedure*) The server selects a subset of clients S_t, $|S_t| = C \cdot K \geq 1$, and distributes the current global model w_t to all clients in S_t. After updating their local models w_t^k to the shared model, $w_t^k \leftarrow w_t$, each client partitions its local data into batches of size *B* and performs *E* epochs of SGD. At the end, the clients upload their trained local models w_{t+1}^k to the central server, which subsequently generates the new global model, w_{t+1} by computing a weighted sum of all received local models. The weighting scheme depends on the number of local training examples, as described through a pseudocode in Algorithm 1, and particularly in equation (3) below [15] [16].

$$w_{t+1} = \sum_{k \in S_t} \frac{n_k}{n_\sigma} w_{t+1}^k \quad (3) \quad \text{where} \quad n_\sigma = \sum_{k \in S_t} n_k$$

ALGORITHM 1: FedAvg

```
Operation on the server side:
initialize $w_0$
for each round t = 0, 1, ... do
    $m \leftarrow \max([C \cdot K], 1)$
    $S_t$ = random set of $m$ clients
    for each client $k \in S_t$ in parallel do
        $w_{i+1}^k \leftarrow$ ClientUpdate($k$, $w_t$)
run equation (3)

Operation on the client side [ClientUpdate(k, w_t)]:
$B \leftarrow$ (split $P_k$ into batches of size $B$)
for each local epoch $i$ from 1 to $E$ do
    for batch $b \in B$ do
        $w \leftarrow w - \eta \nabla \ell(w; b)$
return $w$ to server
```

In Algorithm 1 (FedAvg algorithm) the K clients are indexed by k. B is the local mini-batch size, E is the number of local epochs, and finally η is the learning rate.

Particularly, FedAvg count on the operation of FedSGD. As a typical representation of the FederatedSGD (FedSGD) we could set $C=1$ and then implemented a fixed learning rate of η which depicts to each client k the computation of $g_k = \nabla F_k(w_t)$, representing the average gradient on its local data at the state model w_t, and also the central server aggregates the given gradients and then applies the new $w_{t+1} \leftarrow w_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} g_k$, based on $\sum_{k=1}^{K} \frac{n_k}{n} g_k = \nabla f(w_t)$. Another similar update of the model produced by the following equation, $\forall k,\ w_{i+1}^k \leftarrow w_t - \eta g_k$, which then becomes $w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{i+1}^k$. The last aforementioned equation reflects each client that locally takes one step of gradient descent on the current model with the use of its local data, and continuously the server takes a weighted average of the resulting models. As a result, if the algorithm represented that way, the user be able to add more computation weight to each client by rehearsing the new local data, converted by $w^k \leftarrow w^k - \eta \nabla F_k(w^k)$ a several times before the step of averaging. This approach termed by H. B. McMahan et al. [16] as *Federated Averaging* approach, of better known as *FedAvg*.

14.5.2. CO-OP Algorithm

On the other hand, regarding the asynchronous approach, there is the CO-OP algorithm [20] which we have distinguished. With this approach it is possible to immediately combine any received client model with the global model. Particularly, each client k has an age z_k related with its model and the global model has age z. The model age difference, $z - z_k$, is used to calculate a weight when combing models. This scenario approach is motivated by the fact that in an asynchronous framework, some clients will "*train on outdated models while others will train on more up-to-date*" models [15].

In CO-OP algorithm, a local model will only be combined if $b_l \le z - z_k \le b_u$, for some choice of integers $b_l < b_u$. The intuition behind this rule of common acceptance is that we neither want to merge outdated models $(z - z_k < b_u)$ nor models from overactive clients$(z - z_k < b_l)$. The lower and upper bounds, b_l and b_u can therefore be thought of as an *age filter*. In addition to this, CO-OP inherits all "*hyper-parameters*" from its underlying optimization algorithm, which may be the SGD algorithm, such as FedAvg [15] [27].

The operation of training in CO-OP can be declared as follows: Each client has its own training data, and performs E rounds of an optimization algorithm before requesting the current global model age z from the server. In this aspect, the client has to decide whether or not its age variation meets the restrictions. If the local model is outdated, the client reconciles with the global model and starts over. Otherwise, if the client is active, he simply continues his training. Differently, the local model is uploaded to the server for merging. The pseudocode of CO-OP algorithm is presented in Algorithm 2 [15].

ALGORITHM 2: CO-OP

```
$w = w_1 = ... = w_K \leftarrow w_0$
$z \leftarrow b_l$
$z_1 = ... = z_K \leftarrow 0$
Each client performs $k$ independently runs:
**while** *true* **do**
    Conglomerate a new batch of $B$ samples $D_k$
    $w_k \leftarrow ClientUpdate(w_k)$
    Connection between client – server is ready
    Request and receive the model age $z$ from the server
    **if** $z - z_k < b_u$ **then**
        // Client is outdated
        Fetch $w$, $z$ from the server
        $w_k \leftarrow w$ , $z_k \leftarrow z$
```

```
  else if $z - z_k < b_l$ then
      // Client is overactive
      continue
  else
      // Normal update
      $w_k, z_k \leftarrow UpdateServer(w_{k,} z_k)$ = {
          $w \leftarrow (1-z) \cdot w + z \cdot w_k$ , $z \leftarrow (z - z_k + 1)^{-\frac{1}{2}}$
          $z \leftarrow z + 1$
          return & download w, z
}
```

There are some *age filter restrictions* on CO-OP. Moreover, CO-OP algorithm introduces two additional parameters in its procedures, namely b_l and b_u, but little guidance is provided in order to explain how one should choose these values. Only the intuitive constraint $b_l < b_u$ is given in the original paper that proposes CO-OP [27]. However, arbitrarily choosing these parameters by setting only this limitation in the mind can cause deadlock [15].

If all the clients are considered overactive, thus the algorithm deadlocks. For this reason two additional constraints that should be fulfilled to avoid this deadlock are identified: $b_l < K$ and $b_u < 2b_l$. If the first constraint is unfulfilled, CO-OP is guaranteed to deadlock after *K* updates. This follows from the intuition of b_l; at least b_l normal updates must be performed by distinct clients before a client is allowed another normal update. The second constraint says that a deadlock might occur if the difference between b_l and b_u is too small [15].

14.5.3. Proposed Method

As we can infer, the major advantage of FedAvg algorithm is that orchestrates training through a central server which hosts the shared the global. Additionally, the major advantage of CO-OP algorithm is that makes possible to immediately merge any received client model with the global model. Taking advantage of those two different scenarios we ended up to a scenario that merge these two algorithms in order to have a better efficient model, that selects depending on the occasion if it train the model locally in client or as global in server. Our proposed model named *InFeMo - Integrated Federation Model*.

The "*hyper-parameters*" of our proposal are the same used as the previous models: 1) the fraction of clients C to choose for training, 2) the local mini-batch size *B*, 3) the number of local epochs E, 4) the learning rate η, and 5) the learning rate decay λ. The parameters B, E, η, and λ are typically used when training with SGD, identically with FedAvg and CO-OP models.

In its operation, the IFM algorithm begins with randomly initializing the global model of w_0. Particularly, the operation procedure of the first round of our proposed model demonstrates as: The central-academic server chooses a subset of clients S_i, where will be over 1. Then, the global model which is selected at this time is distributed to all the connected clients S_i. Thereafter, a local model will only be merged if $x_i \le y - y_k \le x_u$, for some choice of integers $x_i < x_u$. The common accepted rule of our scenario is that we want to merge outdated models $(z - z_k < b_u)$. Subsequently, the client updates their local models in order to be shared model, $w_k \leftarrow w$, $y_k \leftarrow y$, each client partitions its local data into batches of size *B* and performs *local updates*. At the end, the clients upload their trained local models w_k and y_k to the central academic server, which subsequently generates the new global model, w_{t+1} by computing a weighted sum of all received local models. The overall weighting scheme is dependent on the number of local training updates, as described through a pseudocode in Algorithm 3, and particularly in equation (3) below.

$$w_{i+1} = \sum_{k \in S_i} \frac{n_k}{n_\sigma} w_{i+1}^k \qquad (4) \qquad \text{where} \quad n_\sigma = \sum_{k \in S_i} n_k$$

Equation (4) exceeds the already defined equation (3).

ALGORITHM 3: Proposed model - InFeMo

```
Client-side operation
$B \leftarrow$ (split $P_k$ into batches of size $B$)
for each local update produced $i$ from 1 to $E$ do
    if ($y - y_k < x_u$) then
        Outdate client Client
        $w_k \leftarrow w$ , $y_k \leftarrow y$
        for batch $b \in B$ do
            $w \leftarrow w - \eta \nabla \ell(w; b)$
    else
        update normally
        $w_k, y_k \leftarrow UpdateServer(k, w_i)$
    return to server $w_k$, $y_k$

Server-side operation
initialize $w_0$
for each round $i$++, ... do
    $m \leftarrow \max(S_t, 1)$
    $S_i$ = random set of $nc$ clients
    for each client $k \in S_i$ in parallel do
        $w_{i+1}^k \leftarrow UpdateClient(k, w_i)$
    run equation (4)
```

Particularly, proposed IFM algorithm keens to provide to the user less waiting time in the queue of the network for each procedure. Due to the decision system of the model the relative data could be decided to be trained locally or globally depending the priority of each occasion. This weighting scheme of the proposed model mainly depends on the number of local updates that could be done in each process.

14.6 Experimental Results

We have made multiple experimental scenarios in order to compare and justify the operation of InFeMo model. Thus, through the experimental scenarios which we have made we have strengthened our suggestion that our proposed architecture is more efficient than the former works. We perform a number of simulations and measurements through which we can realize that we have done a good effort.

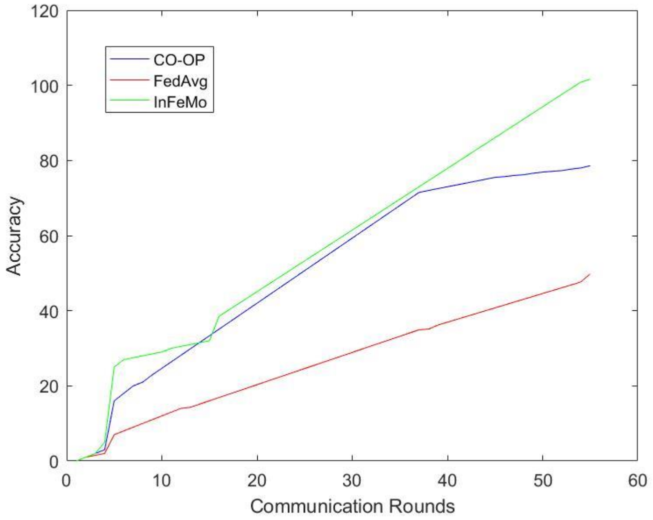


Figure 14.6: Performance comparison of the federated models InFeMo, FedAvg, and CO-OP.

Figure 14.6 describes the better efficient operation provided to the user by applying the InFeMo algorithm in the federated system architecture. As we can observe our proposed model offers more accuracy as long as the communication rounds rises instead of the other two models, the FedAvg and the CO-OP. This means that it could offer a better option of time needed for the user to contact and operate with the academic server. In Figure 14.6, the vertical axis shows the system's accuracy and the horizontal axis shows the communication rounds that have been examined.

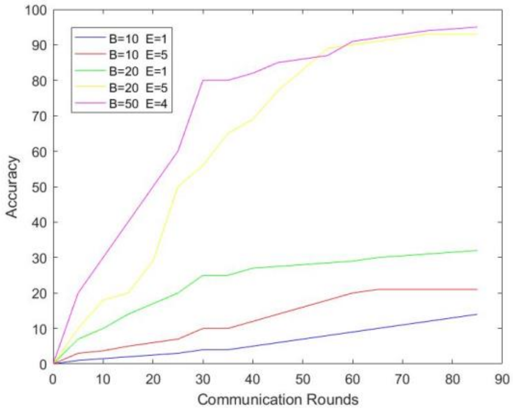


Figure 14.7: Test system's accuracy vs. communication rounds (scenario A).

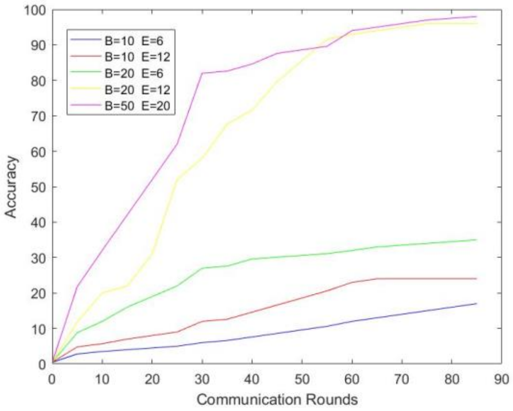


Figure 14.8: Test system's accuracy vs. communication rounds (scenario B).

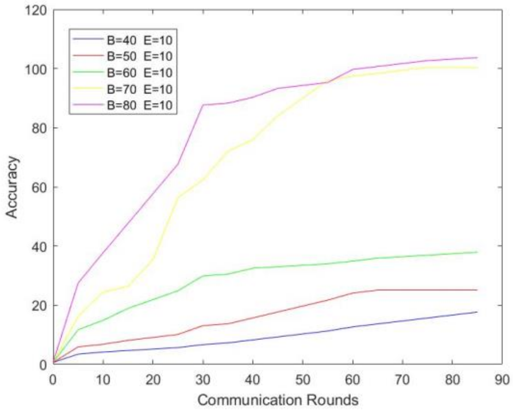


Figure 14.9: Test system's accuracy vs. communication rounds (scenario C).

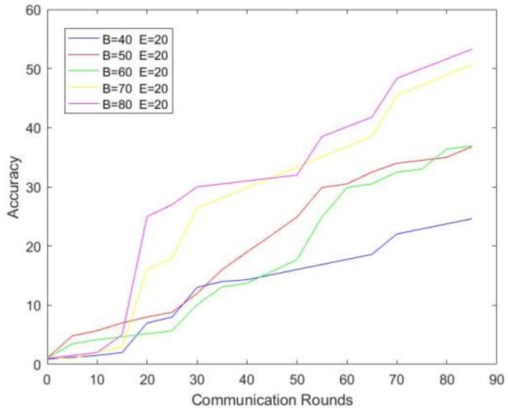


Figure 14.10: Test system's accuracy vs. communication rounds (scenario D).

Figures 14.7, 14.8, 14.9, and 14.10 demonstrate four experimental scenarios that considering the efficiency of different measurements in time. Through these scenarios we can observe that adding more local SGD updates per round could assemble a dramatic reduction in communication costs. The vertical axis shows the system's accuracy and the horizontal axis shows the communication rounds that have been examined for these scenarios. More specific, the expected number of updates per client and per round here is $u=(E[n_k]/B)*E=(n*E)/(K*B)$, where the expectation is over the draw of a random client k. Thus, we can observe that increasing u by varying both E and B is more effective.

14.7 Comparative Analysis

In order to analyze the functionality of our proposed model we have made comparison analysis with some relative previous projects.

The comparative analysis that takes into account here based on two aspects, the architectural model and the CSP-academic server-user communication-authentication model. As regards the architectural model we try to clarify the features of Topology, Encryption Method, Affiliated Technologies, and Cloud Model the works use and include in their function. On the other hand, regarding the CSP-academic server-user communication-authentication model we try to clarify the features of Computation, Authentication, Vulnerability, Trust, and Accessibility of each work compared here.

Work	Topology - Architecture	Encryption method/model	Affiliated Technologies	Cloud Model
Thakur et al. [17]	-	-	-	IaaS
Cai et al. [19]	MemepiC - Traditional Analytics Architecture	-	-	-
Pasquier et al. [20]	Cambridge Flow Control Architecture	IFC	IoT	-
Zhu et al. [21]	Blockchain architecture	Bilinear Pairing Generator	-	SaaS
Yan et al. [22]	-	Attribute-Based Encryption	-	-
Premarathne et al. [23]	-	Security Threat Vulnerability	-	-
Mansha & Kamiran [24]	Mixed topology	-	-	-
Yao et al. [14]	AlexNet Architecture	-	-	-
Wang et al. [25]	Edge Computing Architecture	-	Mobile Edge Computing, IoT	-
Nilsson et al. [15]	Star topology - Artificial neural network architecture	-	-	-
Young et al. [28]	-	-	-	-
McMahan et al. [16]	TensorFlow Architecture	-	-	-
Proposed Model	Mixed Cloud Architecture	AES	Big Data, IoT	SaaS, PaaS & IaaS

Table 14.1: Comparison of architectural model with other former ones

Table 14.1 presents the architecture model characteristics of former related works, compared with our proposed model. The main aspects that studied in order to produce our conclusions are *Topology/Architecture*, *Encryption method or model*, *Affiliated Technologies* integrated in each scenario, and which *Cloud Model* used in each scenario. More specifically, we can observe that most of the works related to Federated Learning Systems (5 of the 6) propose system architecture. Also, regarding the works related to Federated Learning Systems, only one work [25] contributed with another affiliated technology. On the other hand, only the works related to data management in Cloud environment contributed to an Encryption method or model (4 of the 6). This could be resulted because the major goal of this works related to the data, and its usage. Subsequently, through the illustrated findings of Table 1 we can observe that there are not many works in this field that contribute Federated Learning

Systems with another affiliated technologies, and at the same time, proposing a new data management architecture/model.

Work	**Computation**	**Authentication**	**Vulnerability**	**Trust**	**Accessibility**
Thakur et al. [17]			**X**	**X**	
Cai et al. [19]	**X**				**X**
Pasquier et al. [20]		**X**		**X**	**X**
Zhu et al. [21]				**X**	**X**
Yan et al. [22]	**X**			**X**	**X**
Premarathne et al. [23]	**X**	**X**	**X**		**X**
Mansha & Kamiran [24]					**X**
Yao et al. [14]	**X**				
Wang et al. [25]	**X**				
Nilsson et al. [15]	**X**				**X**
Young et al. [28]	**X**				**X**
McMahan et al. [16]	**X**			**X**	**X**
Proposed Model	**X**	**X**	**X**	**X**	**X**

Table 14.2: Comparison of architectural model with other former ones

Table 14.2 lists the basic characteristics studied in this work compared with related previous works analyzed in Section 14.2, which are Computation, Authentication, Vulnerability, Trust, and Accessibility. As we can observe from Table 14.2 the most contributed characteristic is the “Accessibility”, contributed by 9 of 12 works, with most of them contributing the topic “*Big Data management in Cloud*” (5 of 6) works. Additionally, the “Computation” characteristic also contributed most, by 8 of 12 works, with the most works contributed on the topic of “*Federated Learning Scenarios*” (5 of 6 works). Moreover, regarding the characteristics, the “Authentication” is the less contributed characteristic in the related previous works, contributed by 2 of 12, which contributed only from works of the topic “*Big Data management in Cloud*”. Furthermore, the previous related work that contributes the most of the characteristics is U. S. Premarathne et al. work [14], from the topic of “*Big Data management in Cloud*”, which contributes 5 of the 6 characteristics, “Computation”, “Authentication”, “Vulnerability”, and “Accessibility”. Summarizing, the aspects of Table 2 we can observe that arise a gap that our study tries to “fill up” by proposing and presenting a novel system that contributes five major characteristics (“Computation”, “Authentication”, “Vulnerability”, “Trust”, and “Accessibility”) in this field.

Resulting in our findings, as shown by both tables, there is no prior work dealing with integrating specific Cloud models through the federated learning model. Also, none of the earlier work clarifies the encryption model it uses to authenticate users and communicate with the central server. In addition, the proposed model makes grouped and unified use of technologies, as much of the data that is transferred and managed is derived from Internet of Things technology and because of their unique nature, much of the data is characterized as Big Data. In general, there is no mention of consolidated use of technologies in previous work in this field. Further, from the study of the data obtained from the Table 14.2, it seems that very few of the previous papers studied here deal with Authentication and Vulnerability, as well as very few of the previous papers involve all the features listed in Table 14.2 in their study.

On the basis of these data, it seems that the present work is going to fill a scientific gap existing in this field of research. On the one hand, no other architecture model has been studied and proposed so far, which incorporates all Cloud models with a federated scenario, as well as other technologies that may have integrated use with each other. Therefore, we believe that this proposal introduces an innovative idea in the field of federated cloud systems, both in the field of administration, as well as end-user communication with the server, which is also related to security and immediacy.

14.7.1. Distributed Cloud vs. Federated Cloud

In this subsection will present the main differences between Distributed and Federated technique. These key differences also extend to their use in Cloud environments.

The fundamental dissimilarity among federated learning and distributed learning counts on the pretensions made on the features of the local datasets [29] [30]. On the one hand the distributed learning inventively targets at parallelizing computing power while on the other federated learning inventively targets at training on heterogeneous datasets. Additionally, distributed learning targets at training a single model on innumerable servers, a common underlying hypothesis is that the local datasets are i.i.d. and roughly have the same size. Federated learning does not count on speculations such these, rather the datasets are typically heterogeneous and their sizes might span various orders of magnitude [31].

Moreover, through the federated technique a problem that arises in the distributed technique could be solved. The problem is: "*Given a set of overlapping distributed queries bound to perform multiple aggregation operations on a given set of data sources, place the aggregation operators within the communication network to minimize the cost of data movement across the communication edges of network*" as previously set by S. Mansha & F. Kamiran [24]. This issue was solved by S. Mansha & F. Kamiran [24] in their work use an evolutionary algorithm that expands a federated learnig technique.

Also, regarding to former works in the field, such as those of A. Nilsson et al. [15], H. B. McMahan et al. [26], and J. Konecny et al. [32], typically considered that the distributed optimization algorithms achieve that:

- ✓ Data is regularly distributed over clients
- ✓ Client-side data are independent and identically distributed (widely known as i.i.d.) illustrations from the overall distribution
- ✓ The number of clients is much smaller than the average number of locally available training examples per client

As a result, the distributed data center optimization typically obligates control over the data distribution since these approaches count on balanced and i.i.d. data speculations [15]. In the other hand, the federated learning proposes to have a number of edge devices perform its procedure tasks locally and as a result only communicate an updated model to a collaborating server with them [15]. Also, count on the novel law commitments, federated learning utilizes the General Data Protection Regulation's (GDPR) data minimization principle [33] since only the learned model, and no raw data, is produced centrally [15].

14.8 Chapter Summary

Cloud Computing could be used to be a base technology for many technologies due to its type of services. Cloud Computing provides new generation of services which aims to offer accessibility to information, applications and data from any place at any time. Moreover, this work presented and described a new system architecture based on Cloud Computing, and count on the novel scenario of Federated Learning, which called Integrated Federated Model - InFeMo. Our model incorporates all Cloud models with a federated learning scenario, as well as other technologies that may have integrated use with each other. The major motivation of InFeMo is to offer provide a more efficient system architecture and environment for the academic users with the aim to data management. This efficiency of our proposal counts on its operation, because it decreases the number of rounds of communication that needed to train a scenario model by using a federated Cloud system, and as a result it makes the user that uses this system to wait less. System's federated algorithm relies on the advantages of the former models of FedAvg and CO-OP algorithms. Consequently, we ended up to our new scenario that merges these two algorithms aiming to have a more efficient model, which selects the training model depending on each occasion.

As a result, due to our work tries to fill a scientific gap in the field of federated cloud systems, we can make more researches and experiments in order to achieve and explore the new opportunities arising in this field of study. Based on our research and the comparative analysis, no other architecture model has been studied and proposed so far, which incorporates all Cloud models with a federated scenario, as well as other technologies that may have integrated use with each other. So, we keen on to work on this field trying to find out new aspects that could lead us to efficient and more secure communication between the user and the central server. Also, we could try to involve an IoT scenario of sensors and a Smart Building scenario in order to find out new methods and aspects that arise here.

There are also other areas where this proposed model could be applied, beyond the academic community, offering multiple benefits. As already mentioned above, the academic community can offer users a more efficient environment, offering less waiting time and ease in the mass management of their data. Regarding the health sector, where the proposed model could also be applied, as it would mainly facilitate the medical staff in the easier access to the sensitive medical data and their analysis, which can be performed in the specific Cloud environment of this system, more immediate and efficient. Thus, for example, the doctor will be able to receive the data needed through the smart phone directly, and also due to the use of the federated scenario to enable data analysis applications in order to provide data immediately and quickly. Another area of application of the proposed model could be industry. There it would facilitate the most efficient supervision of the production process of a production chain. Based on the federated scenario of the proposed model, the managers of each related part of the production will be able to receive data analysis components, but also to have faster and more direct access to them. In addition, even in the administrative part of an industry it could help in better and faster access to data, by giving the users the ease of using a Cloud infrastructure without the necessary choice of a specific provider. These could be the next areas of the future continuation of our current research.

14.9 Chapter References

[14] X. Yao, C. Huang, L. Sun, “Two-Stream Federated Learning: Reduce the Communication Costs”, in Proceedings of 2018 IEEE Visual Communications and Image Processing (VCIP), 9-12 December 2018, Taichung, Taiwan, Taiwan. [DOI: 10.1109/VCIP.2018.8698609]

[15] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, M. Jirstrand, “A Performance Evaluation of Federated Learning Algorithms”, in Proceedings of DIDL '18: Proceedings of the Second Workshop on Distributed Infrastructures for Deep Learning, December 2018, pp. 1-8, Middleware '18: 19th International Middleware Conference Rennes France. [DOI: 10.1145/3286490.3286559]

[16] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y Arcas, “Communication-Efficient Learning of Deep Networks from Decentralized Data”, in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, JMLR: W&CP, volume 54, 20-22 April 2017, Fort Lauderdale, Florida, USA. [arXiv:1602.05629]

[17] R. Shokri, V. Shmatikov, “Privacy-preserving deep learning”, in Proceedings of 53rd Annual Allerton Conference on Communication, Control, and Computing

(Allerton), 30 September – 2 October 2015, Allerton Park and Conference Center, USA.

[18] S. Thakur, J. G. Breslin, "A Robust Reputation Management Mechanism in the Federated Cloud", IEEE Transactions on Cloud Computing, vol. 7, issue: 3, pp. 625-637, July-September 2019. [DOI: 10.1109/TCC.2017.2689020]

[19] Q. Cai, H. Zhang, W. Guo, G. Chen, B. Chin Ooi, K.-L. Tan, W.-F. Wong, "MemepiC: Towards a Unified In-Memory Big Data Management System", IEEE Transactions on Big Data, vol. 5, issue: 1, pp. 4-17, March 2019. [DOI: 10.1109/TBDATA.2017.2789286]

[20] T. F. J.-M. Pasquier, J. Singh, D. Eyers, J. Bacon, "CamFlow: Managed Data-sharing for Cloud Services", IEEE Transactions on Cloud Computing, vol. 5, issue: 3, pp. 472 - 484, July-September 2017. [DOI: 10.1109/TCC.2015.2489211]

[21] L. Zhu, Y. Wu, K. Gai, K.-K. R. Choo, "Controllable and trustworthy blockchain-based Cloud data management", Elsevier, Future Generation Computer Systems, vol. 91, pp. 527-535, February 2019. [DOI: 10.1016/j.future.2018.09.019]

[22] Z. Yan, L. Zhang, W. DING, Q. Zheng, "Heterogeneous Data Storage Management with Deduplication in Cloud Computing", IEEE Transactions on Big Data, vol. 5, Issue: 3, pp. 393-407, September 2019. [DOI: 10.1109/TBDATA.2017.2701352]

[23] U. S. Premarathne, I. Khalil, Z. Tari, A. Zomaya, "Cloud-Based Utility Service Framework for Trust Negotiations Using Federated Identity Management", IEEE Transactions on Cloud Computing, vol. 5, Issue: 2, pp. 290-302, April-June 2017. [DOI: 10.1109/TCC.2015.2404816]

[24] S. Mansha, F. Kamiran, "Multi-Query Optimization in Federated Databases using Evolutionary Algorithm", in Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications, 9-11 December 2015, Miami, FL, USA. [DOI: 10.1109/ICMLA.2015.125]

[25] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, K. Chan, "Adaptive Federated Learning in Resource Constrained Edge Computing Systems", IEEE Journal on Selected Areas in Communications, ver. 99, pp. 1-1, March 2019. [DOI: 10.1109/JSAC.2019.2904348]

[26] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, et al. 2016. Communication-efficient learing of deep networks from decentralized data. arXiv: 1602.05629

[27] Yushi Wang. 2017. CO-OP: Cooperative Machine Learning from Mobile Devices. Master’s thesis. Dept. Elect. And Comput. Eng., Univ. Alberta, Edmonton, Canada.

[28] B. Young, R. Bhatnagar, G. Tatavarty, H. Bian, “Covariance Matrix Computations with Federated Databases”, in Proceedings of ICMLA '07: Proceedings of the Sixth International Conference on Machine Learning and Applications, 13-15 December 2007, pp. 172-177, Cincinnati, OH, USA. [DOI: 10.1109/ICMLA.2007.36]

[29] J. Konecny, H. B. McMahan, D. Ramage, “Federated Optimization:Distributed Optimization Beyond the Datacenter”, ArXiv, pp. 1-38, November, 2015.[arXiv:1511.03575] [Retrieved March 2020] [link: https://arxiv.org/abs/1511.03575]

[30] J. Pei, P. Hong, K. Xue, D. Li, “Efficiently Embedding Service Function Chains with Dynamic Virtual Network Function Placement in Geo-Distributed Cloud System”, IEEE Transactions on Parallel and Distributed Systems, vol. 30, issue: 10, pp. 2179 – 2192, October 2019. [DOI: 10.1109/TPDS.2018.2880992]

[31] K.-Y. Chen, Y. Xu, K. Xi, H. J. Chao, “Intelligent virtual machine placement for cost efficiency in geo-distributed cloud systems”, in Proceedings of 2013 IEEE International Conference on Communications (ICC), 9-13 June 2013, Budapest, Hungary. [DOI: 10.1109/ICC.2013.6655092]

[32] Jakub Konecny, H. Brendan McMahan, Daniel Ramage, and Peter Richtarik. 2016. Federated Optimizetion: Distributed Machine Learning for On-Device Intelligence. arXiv: 1610.02527

[33] European Commission. 2018. What data can we process and under which conditions? Retrieved 14 March 2020, from https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr/what-data-can-we-process-and-under-which-conditions_en

Chapter 15

EEIBDM: A Reinforcement & Federated Learning scenario for Efficient Industrial IoT-based Big Data Management in Cloud

Paper 28 - EEIBDM: A Reinforcement & Federated Learning scenario for Efficient Industrial IoT-based Big Data Management in Cloud

Paper titled "EEIBDM: A Reinforcement & Federated Learning scenario for Efficient Industrial IoT-based Big Data Management in Cloud" authored by C. L. Stergiou and K. E. Psannis has been submitted in IEEE Transactions on Sustainable Computing, and it is under review procedure.

15.1 Introduction

This work makes an effort to survey the multiple open challenges and issues in the field of energy-efficient industrial IoT-based Big Data management in Cloud environments. Aspects and challenges arise from the fields of Artificial Intelligence scenarios of the Cloud infrastructures, Artificial Intelligence techniques of Big Data Analytics in the Cloud environments and Federated Learning Cloud systems try to be clarified. Additionally, Reinforcement Learning is a novel technique that allows large data centers such as Cloud data centers to affect a more energy-efficient resource allocation. Moreover, take into account all the works that have been done in the field, I propose an architecture that tries to combine the features offered by the several Cloud Providers to emerge and achieve an Energy-Efficient industrial IoT-based Big Data Management Framework (EEIBDM) established outside of every user, in Cloud environment. IoT data could be integrated with techniques such as Reinforcement and Federated Learning to achieve a Digital Twin scenario, by creating novel digital simulation models. Furthermore, I propose an algorithm for delivering the energy consumption of CPU through the evaluation of EEIBDM framework. Finally, some future directions as an expansion of my research are illustrated.

15.2 Related Work

For the purpose of our work we study and analyze former literature of the field of related topics. The following paragraphs illustrate the works which contributed significantly in our study.

15.2.1 Efficient Big Data Management in the Cloud

Management in Cloud environments. Thus, for the purpose of our research we have studied and analyzed previous literature researches that have been made in the field of Efficient Big Data management in Cloud environment [7] [8] [9]. The following paragraphs present the previous to our study research papers.

Aujla and Kumar [7] present MEnSuS which is an efficient system for energy management with sustainability of Cloud Data Centers in Edge-Cloud Environment with the use of SDN. The proposed scheme is a support vector machine-based workload classification approach. Al-Dulaimy et al. [8] investigate the design and implementation of virtual machine management strategies for energy efficient cloud data centers, and also they propose a novel model in order to solve the problem of VM placement. Khan et al. [9] propose a wide range of heuristic and meta-heuristic VMC algorithms, because of VMC is a NP-hard problem, which aims to achieve near-optimality. Additionally, they classify and critically review VMC algorithms from large number of viewpoints so that the readers can be truly assessed.

15.2.2 AI Cloud Scenarios

Furthermore, there are a number of remarkable works associated with the novel scenarios of Artificial Intelligence Cloud systems regarding the use of data analytics [10] [11] [12]. The following paragraphs present the relative to our study research papers.

Weber et al. [10] consider the particular needs of users of cloud computing resources, wishing to manage the resources, and they try to present an approach to rollback for cloud management, that wrapper the Cloud management API, and uses AI Planning techniques to find an appropriate undo sequence. Brown and Kauchak [11] talk about and share novel educational approximations that teach or leverage Artificial Intelligence and its many subsections, including robotics, machine learning, natural language processing, computer vision, and others at all layers of education. Rad et al. [12] present Cloud-eLab platform, which is an open and interactive Cloud-based learning platform for AI Thinking, intending to infuse two aspects: i) Deep and Wide learning, ii) Cognitive and Adaptation learning notions for education.

15.2.3 AI Big Data Analytics in the Cloud

Moreover, several remarkable works associated specific with Artificial Intelligence systems of Big Data Analytics in Cloud environments [13] [14] [15] [17] [18]. The following paragraphs present the relative to our study research papers.

Wu et al. [13] review historical perspectives of the term "*Big Data*" and the correlated analytics show that Big Data is not only 3Vs, but it could be divided into 32 Vs. Specifically, 9 Vs covering the cardinal motivation inside the Big Data term, which is to merge Business Intelligence (BI) count on various hypothesis or statistical models so that Big Data Analytics (BDA) could assist decision makers to make useful predictions for some fatal decisions or researching results. Ahmed et al. [14] make an investigation of the recent advances in BDA for IoT systems like the key requirements for managing Big Data and for activating analytics in an IoT environment. Lee et al. [15] consider the 5C architecture proposed by a previous work of Lee et al. [16], and propose an insight into the ongoing of AI technologies and the eco-system needed to harness the power of AI in industrial applications. Wan et al. [17] present a vertically-integrated four-tier CaSF (cloud-assisted smart factory) architecture. With this proposal Wan et al. aim to highlight the role and contingent of Cloud Computing and AI in meliorating the smart factories' performances, like system flexibility, efficiency, and intelligence, they completely summarize and explain the AI application in a cloud-assisted smart factory (CaSF). Khan et al. [18] explore the current research, challenges, open problems, and future research direction for the several problems need to be confronted and risks that need to be moderated before practical applications of this synergistic model which can be popularly used.

15.3 Proposed Approach

Based on previous works, with the aim to succeed a new efficient system for industrial IoT-based Big Data Management in Cloud Environment, count on Reinforcement and Federated Learning techniques, we formulate and design the architecture of our proposed system.

Industrial IoT-based data management in Cloud infrastructure became popular the recent years due to the help of software 'infrastructure' which supports efficiently the operation of data centers and furthermore the Cloud data centers. Due to the maximum need of hardware infrastructure and the need of the continuously update the data centers providers intend to host all their infrastructure in data centers that can support as many as they can customers, thus adopt virtualization. Virtualization is a new technique where users are granted virtual platforms, rather than physical ones aiming to resolve several operational and maintenance issue in data center. Also, virtualization is an effective way to offer the solution of management of dynamic resources on Cloud environment.

Through virtualization we can use Virtual Machines (VMs). VM is a number of identical, isolated execution environments on a single computer, each of which emulates the host computer. As a result, this gives the user the illusion of having his/her own physical machine. These VMs can be used through emulators-simulators. An emulator used to simulate a hardware platform, generally in order to allow running multiple Operating Systems simultaneously, and support foreign code on a given platform. The emulator that we used in our research is CloudSim, which operates on Eclipse.

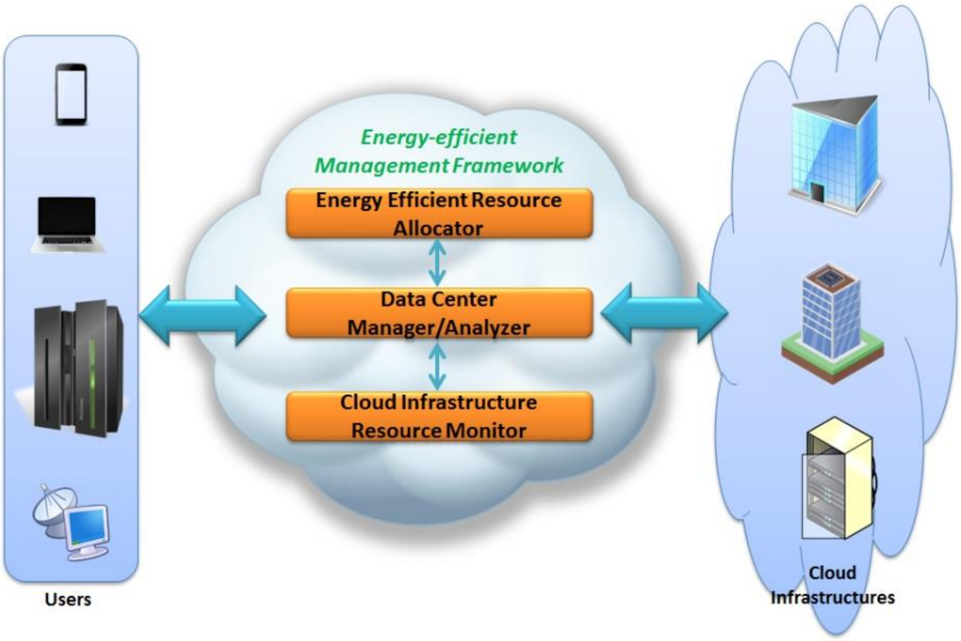

Energy-efficient Management Framework
Energy Efficient Resource Allocator
Data Center Manager/Analyzer
Cloud Infrastructure Resource Monitor
Users
Cloud Infrastructures

Figure 15.1: System architecture.

The proposed architecture scenario tries to combine the features offered by the several Cloud Providers in order to emerge and achieve an Energy-Efficient industrial IoT-based Big Data Management Framework (EEIBDM) established outside of every user, in the Cloud. The proposed system architecture, count on the IaaS and PaaS models of the Cloud offered by the cooperative Cloud Providers attempt to meet particularly metrics such as CPU resources, memory amount, storage availability, and system's performance in terms of execution time. As we can observe from Figure 15.1, each different type of user could use the Cloud Infrastructures of the various cooperative Cloud Providers through the EEIBDM, which offers three major advantages: 1) Energy Efficient Resource Allocator, 2) Data Center Manager/Analyzer, 3) Cloud Infrastructure Resource Monitor.

15.3.1 Energy-Efficient resource allocation with Reinforcement Learning

Reinforcement Learning (RL) could be defined as "*a process of learning by interactions with dynamic environment, which generates the optimal control policy for a given set of states without requiring domain knowledge of the environment*" regarding our literature study [19] [20]. Several functions that endless affect the extracted reward through the learning process of the system. Furthermore, RL count on two fundamental operations: 1) trial-and-error search, and 2) delayed reward.

Thus, the problem formulation of the energy-efficient resource allocation in a Cloud environment, which aims to be addressed with Reinforcement Learning, can be introduced as a sort of optimization problem in terms of the already wide known *Markov Decision Processes*. Our research goal lays in the formulation of the aspects of the resource allocation issue, while interacting with the Cloud environment aiming to achieve an optimal decision.

15.3.2 Federated Cloud System Modeling

We try to model a Cloud Coordinator entity as a requirement for federate multiple Clouds. Cloud Coordinator (ClCo) is responsible for monitoring and managing the internal state of a data center entity, except of communication with other data centers and end-users in the simulation environment. The produced information received from ClCo as a part of the monitoring process, which is on through the simulation period that utilized for making decisions associated to inter-cloud provisioning. CloudCoordinator's functionality might be defined as similar to the functionality offered by large businesses. Resulting this, when an engineer of a physical Cloud system demands to federate services from several Cloud Providers, it will be required a development of a CloudCoordinator. So, aspects associated to communication and negotiations to ex-entities are isolated from the core of the data center, in order to have an entity able to manage the federation scenario of Cloud data centers. Consequently, CloudSim operation provides to every cloud developer to speed up the use of application service performing tests through an entity such as ClCo.

15.4 Evaluation of Proposed System

The evaluation of our proposed work presented in this section.

15.4.1 Energy Efficiency on Data Center

It is a fact that the industrial large data center infrastructures assume CPs huge energy consumption cost of the resources of the infrastructures, which lead to considerable raise of environmental costs. This is a major issue in order to keen us to the energy cost and carbon footprint of Cloud systems. Aiming to minimize the energy consumption, intelligent mechanisms need to be built with the ability to be managed across different heterogeneous machines.

Regarding our previous study on literature review, to achieve energy efficiency we have to integrate mechanisms of Reinforcement Learning and Federated Learning, with the aim to reach the ability of a Cloud system has to decrease the consumption of the resources which are not in use.

The widely accepted unit of measurement for Energy Efficiency in a data center is Power Usage Effectiveness (PUE). PUE is already defined by Armbrust et al. [21] as a green grid component, and also it is states the ratio of the total amount of the power used in a data center facility per the power which delivered to the IT hardware equipment. Specifically, PUE could be delivered by the following equation:

$$PUE = \frac{TFP}{ITEP} \qquad (1)$$

In equation (1) the value of TFP represents the Total Facility Power which demonstrates the data centers' entirely power that is delivered. On the other hand, ITEP represents the IT Equipment Power which demonstrates the energy facilities consumed from the equipment which is used in order to manage, process, transfer, store, operate, and route data through the data. Due to the experiment analyzed before by Koutitas and Demestichas [22], the result of equation (1) emphasizes mostly on energy consumption of Cloud data center's IT equipment which has the 30% of the whole data center. As a result, the PUE of the equation (1) can be configured as follows:

$$PUE = \frac{NIT_{pc} + (CPU_{pc} + NonCPU_{pc})}{(CPU_{pc} + NonCPU_{pc})} \qquad (2)$$

In the case of equation (2), in order to produce the value of PUE, we have to calculate the sum between the value of NITpc, that represents the NonIT Equipment power consumption (30% of the total), and the summary of CPUpc, represents CPU power consumption (40% of the total), and NonCPUpc, represents NonCPU power consumption (10% of the total), divided by the summary of CPUpc and NonCPUpc. The energy consumption of CPU could be delivered from our proposed method presented in Algorithm 1, and represents the use of the resource allocation scenario provided by our Energy-Efficient Big Data Management (EEBDM) framework. Last but not least, the overall-high value of power consumption of the CPU in data center produces the high expenditure of cooling system, but in our scenario it is not necessary.

15.4.2 EEIBDM's framework resource allocation evaluation

The Algorithm 1 introduced as a novel resource allocation algorithm. This algorithm embedded and tested in the software of CloudSim toolkit. All the aspects of our proposed framework are included as the part of extensive heuristic in CloudSim toolkit. As already mentioned in section 3, CloudSim, as virtualization software, and through the literature could be consisted as a scalable simulation framework that enables innovative support for modeling, simulation, and experimentation of virtualized data centers in Cloud environments, and in addition Cloud management services for all the components such as VMs, memory, storage, and bandwidth, under various capabilities, configurations, and domains. Finally, CloudSim could support characteristics that models and simulates environments based on large-scale Cloud, resource allocation policies of energy efficient scenarios, service brokers, virtualization techniques, federated Cloud systems of CPs, and established network connections.

Algorithm 1: EEIBDM's framework resource allocation scenario

As inputs of the method accepts:

- ✓ The number of the different hosts operates in the data center initialized: *NoHost*
- ✓ The number of the various VMs operates in the data center initialized: *NoVM*
- ✓ The CPU's workload value counts on the different users of the system per second: *cpuw*
- ✓ The discount factor of the system: *dfs*
- ✓ The particular upper limit of the learning process: *U*

The method exports as output:

- ✓ optimized distribution of the used VMs: *overall allocation*

Proposed Method:

```
initialize Host(NoHost) // create Hosts operating in data center with specific features
initialize VM(NoVM) // create VMs operating in data center with specific features
initialize CPUW(cpuw) // initialize CPU's workload and data center components
create Environment() // set up state set S, action set A and initialize K values and F values
for VM ϵ NoVM // each VM contained in Number of VMs
    for Host ϵ NoHost // each Host contained in Number of Hosts
        $S_b$ = {edc, h, vm} // convey values of energy of data center, host and Vm of the specific state $S_b$
        for i, i=0,1,2,3,…,U
```

```
                $A_i = A \in \max_A * K_i(S_i, A')$
                with $A_i$ count $A_{i+1}$
                recompense $F_{i+1}$
                $K_{i+1}(S_i, A_i) \leftarrow F + [dfs \cdot \max_{A^F} \cdot K_i(S_{i+1}, A)]$
                // update the existing value
                $S_i = S_{i+1}$ // distribute the next host
            **end**
        **end**
        **return** h
        distribute(Host, VM) // allocate new host and VM
    **end**
    **return** overall allocation
```

Taking into account the characteristics provided by our proposed method, CloudSim is able to exploit novel construct heuristics in order to assess the performance obstacles associate to the service delivery and provisioning policies in resource management techniques. Therefore, the existing architecture of CloudSim software supports Cloud infrastructure service management, but unfortunately does not care either about the energy consumption of a data center, or the PUE value.

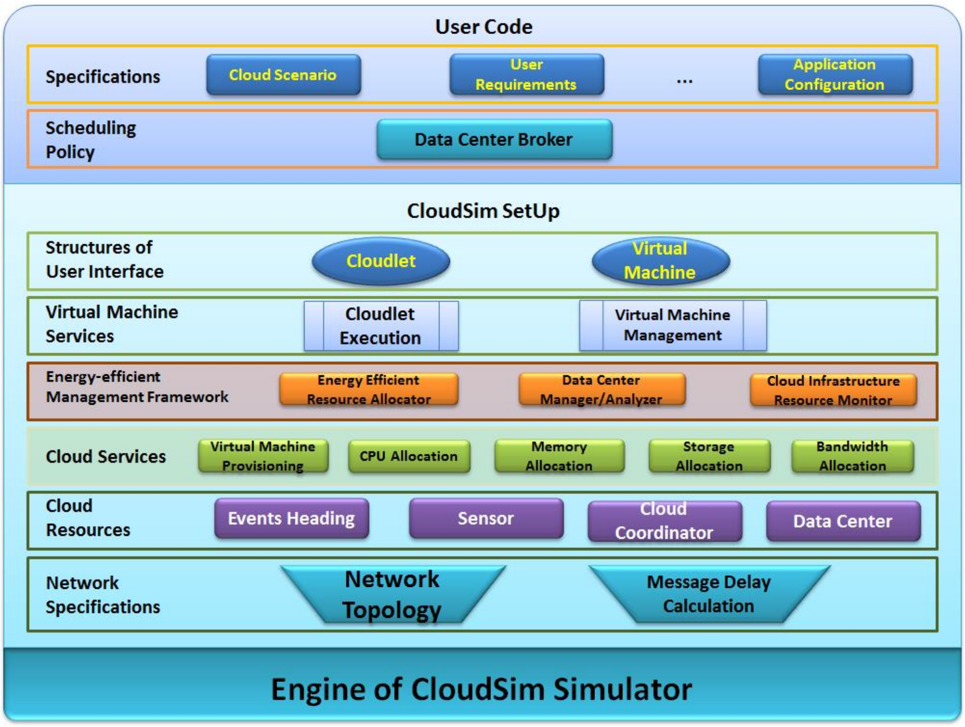


Figure 15.2: CloudSim architecture emerges with EEIBDM.

Figure 15.2 illustrates the existing architecture of CloudSim software integrated with our proposed EEIBDM state in order to achieve an energy-efficient resource allocation service through the existing architecture. The EEIBDM framework, as we can observe from figure 15.2, placed in the middle of the core of CloudSim set-up component.

15.4.3 Reinforcement based Cloud evaluation resource allocation

Based on the literature study we know that a Service Level Agreement (SLA) manager probes the utilization of the CPU related to all the hosts conducted in the data center abbot to VM allocation of a Cloud environments aiming to secure the covenant allocation of SLA metrics. This calculation showed in equation (4) below:

$$EC_n = \int_{t0}^{t1} F(u(t))dt \qquad n = 1, 2, ..., k \qquad (4)$$

Equation (4) represents the value of overall energy consumption of a specific host that is functioned of the overall calculation time, illustrated to EC_n. Moreover, CPU utilization corresponds to *u(t)*, the period of overall calculation time for each host defined as n, and its range defined from the total number of hosts contribute to the data center starting from 1 to *k*.

$$EC = \sum_{n=1}^{k} EC_n \qquad (5)$$

Resulting equation (4) the calculation of the total amount of energy consumption of the data center referred to the energy consumption of all the contributed hosts in the data center could be better described be the previous equation (5).

Data Center (Host)	Virtual Machine
12GB RAM memory	512MB RAM memory
2TB storage memory	20GB storage memory
2 x CPU with 1000 MPIS capacity	1 x CPU with 1000 MPIS capacity
Time-shared VM scheduler	Time-shared Cloudlet scheduler

Table 15.1: CloudSim Configuration – Reinforcement Cloud-based evaluation

Table 15.1 shows that in this scenario the hardware set-up of each Cloud data center serves structured.

15.4.4 Experimental Results of simulating Reinforcement Cloud Evaluation

In this subsection, the experimental results of the Reinforcement Cloud simulation in CloudSim demonstrated and analyzed.

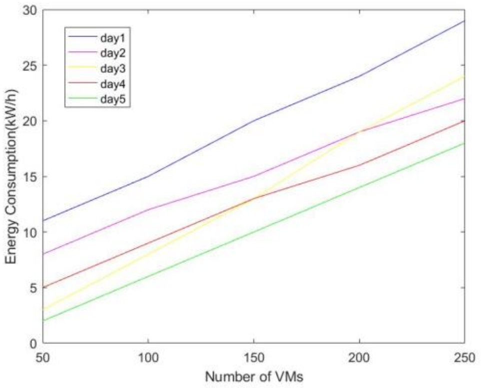


Figure 15.3: Energy consumption during 5 days simulation of Reinforcement Cloud in CloudSim.

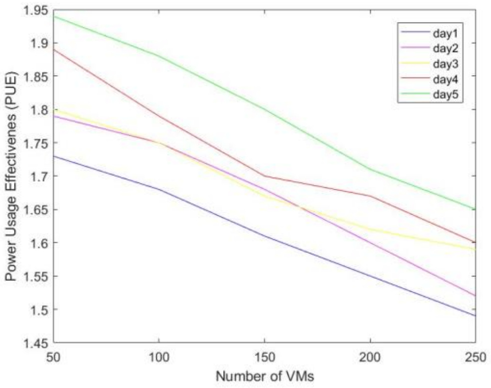


Figure 15.4: Power Usage Effectiveness during 5 days simulation of Reinforcement Cloud in CloudSim.

Figure 15.3 demonstrates the effects of energy consumption (EC) of the simulation to the need of VMs of each day proceeds. More specific, it is shown how the energy of the data center consumed by serving the requests from VMs in the 5 days schedule, for five states of the used number of VMs. Assuming figure 15.3, as the number of VMs increases, the value of energy consumption in the data center increases. Also, as an overall statistic, we can state that for 50 VMs the EC is bellow 12kW/h, for 100 VMs is bellow 15kW/h, for 150 VMs is under 20kW/h, for 200 VMs is under 25kW/h, and finally for 250 VMs is under 29kW/h. Moreover, we can conclude that during the days past the EC decreases.

Figure 15.4 shows that the value of PUE is a decreasing value as the need to use more VMs grows. Thus, the major goal of achieving an energy efficient use when more VMs required reached. Count on figure 15.4 when there is a need of 50 VMs the PUE is under 1.95 for all 5 days. Additionally, for 100 VMs the PUE is below 1.89, for 150 VMs is below 1.80, for 200 VMs is under 1.70, and fro 250 VMs is not more than 1.65. According to the literature range of an energy efficient value of PUE, is between 1 and 2. Consequently, the proposed algorithm and method achieves the energy efficiency level of a data center regarding the results demonstrated in figure 15.4.

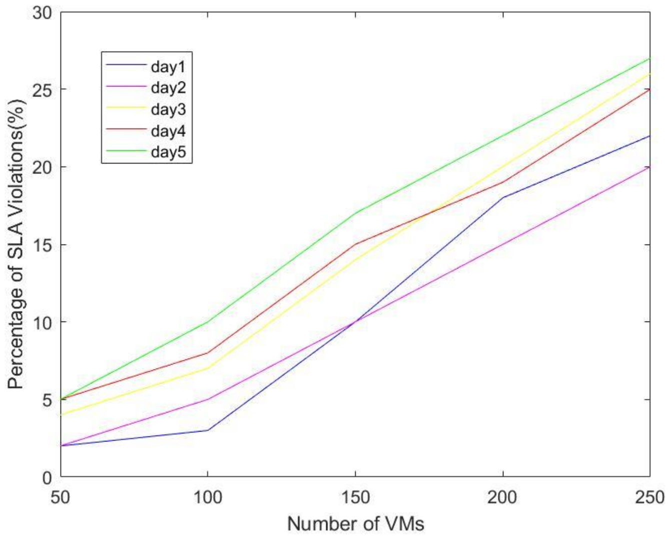


Figure 15.5: Percentage of SLA violation during 5 days simulation of Reinforcement Cloud in CloudSim.

Figure 15.5 shows the percentage value of SLA violation performed during 5 days simulation of Reinforcement Cloud in CloudSim. It is demonstrated that the more the number of VMs increased the more the percentage of SLA violations increased, so they are analogous values. An increase of about 5% in SLA violation resulted for 50 VMs allocation, while an increase of 10% in SLA violation resulted for 100 VMs allocation, 16% for 150 VMs allocation, 20% for 200 VMs allocation, and 27% for 250 VMs allocation.

15.4.5 Federated Cloud evaluation features

In order to produce a more efficient system we model and simulate a federated Cloud network in CloudSim. Thus, we model a system of three CPs federations and a connection to a User Broker. Each CP institutes a sensor, which has the responsibility to sense the availability of information associated to the data center hosts dynamically. Subsequently, the measures of this sensor delivered to the ClCo where the produced information utilized in undertaking load-migration decisions. As a result, in this system would be clearly performed transmigration of the available VMs through the cooperative CPs taking into account the possibility of the initial CP is not able to provide the requested number of available VM slots. The topology of this scenario represented in figure 15.6, where demonstrated the Cloud Providers federation.

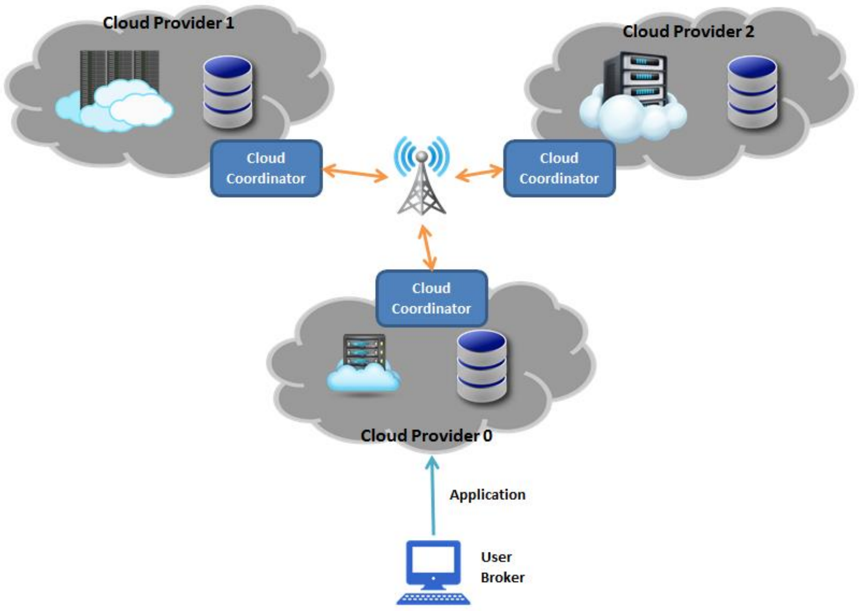


Figure 15.6: Federated Cloud data centers topology.

In the aforementioned scenario the model components of CloudSim simulation represented in Table 15.2. The performance results of federated Cloud simulation in CloudSim shown in Table 15.3.

Data Center (Host)	Virtual Machine (x50)
100 computing hosts	1 x VM = 1 x Cloudlet
12GB RAM memory	512MB RAM memory
2TB storage memory	20GB storage memory
2 x CPU with 1000 MPIS capacity	1 x CPU
Time-shared VM scheduler	Time-shared Cloudlet scheduler
Cloudlet length is 18000000 MIs	

Table 15.2: CloudSim Configuration – Federated Cloud set-up

Performance metrics	Federation results	Non-Federation results	
Average turn around time (sec)		4241.45	8782.9
Makespan (sec)	7653.62	14609.41	

Table 15.3: CloudSim Performance results of federated Cloud

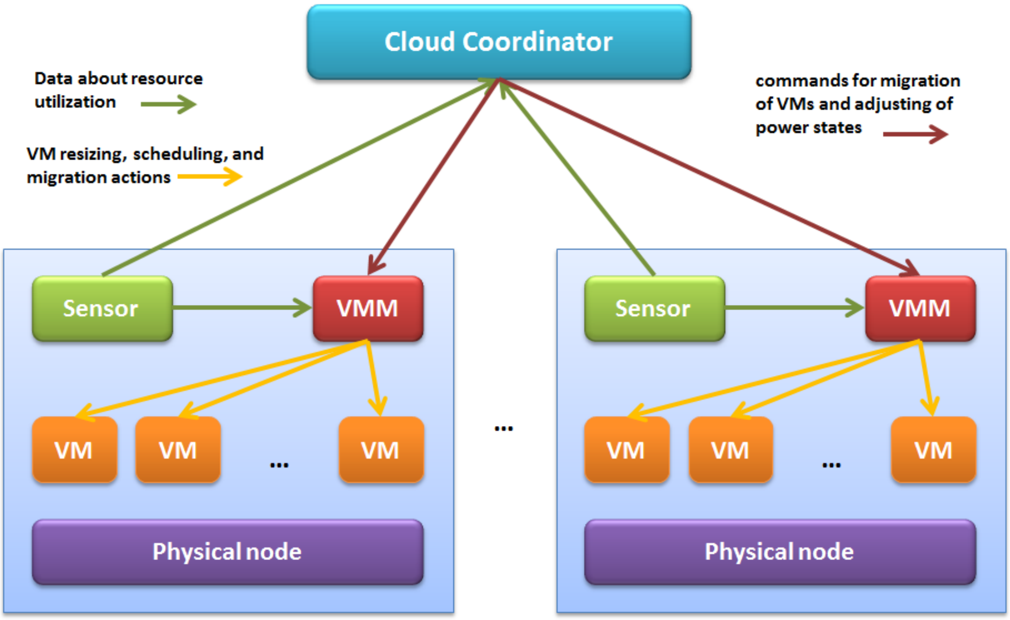

Cloud Coordinator
Data about resource utilization
VM resizing, scheduling, and migration actions
commands for migration of VMs and adjusting of power states
Sensor
VMM
VM
VM
...
VM
Physical node
...
Sensor
VMM
VM
VM
...
VM
Physical node

Figure 15.7: Energy-Conscious management architecture.

Assuming the operation of ClCo, and the previous proposed methods, the system's Cloud Computing architecture could be represented in figure 15.7. Figure 15.7 considers EEBDM system included techniques relay on data center, ClCo and Sensor components. Through the embedded sensors the ClCo would be able to monitor during the time the performance of each active VM. Thus, the VMM gets the real-time data, and then use this data in order to perform particular resize of the VMs needed. Finally, ClCo performs allocation of the VMs applying VM migration and additionally changes the power state of each node, following the rules of resources utilization.

Work	NC	BPT	EES	AIT	AIA	AII	AIC	RM	VC	RC	IS	IA	BA	BM	FM	RM	P	A	F
Aujla & Kumar **[9]**	**M**	**L**	**H**	**L**	**L**	**L**	**L**	**M**	**-**	**X**	**L**	**M**	**L**	**H**	**-**	**-**	**X**	**-**	**-**
Al-Dulaimy et al. **[10]**	**L**	**L**	**H**	**L**	**M**	**L**	**M**	**H**	**X**	**X**	**M**	**L**	**L**	**H**	**-**	**-**	**X**	**X**	**-**
Khan et al. **[12]**	**M**	**H**	**M**	**L**	**L**	**L**	**H**	**H**	**X**	**-**	**M**	**L**	**M**	**M**	**-**	**X**	**-**	**-**	**X**
Weber et al. **[13]**	**L**	**H**	**L**	**M**	**H**	**L**	**H**	**L**	**X**	**X**	**M**	**L**	**L**	**H**	**X**	**-**	**X**	**-**	**-**
Brown & Kauchak **[14]**	**M**	**M**	**L**	**H**	**H**	**H**	**H**	**M**	**-**	**X**	**M**	**M**	**H**	**H**	**X**	**-**	**X**	**-**	**X**
Rad et al. **[15]**	**M**	**M**	**L**	**H**	**H**	**M**	**H**	**M**	**X**	**-**	**H**	**L**	**L**	**L**	**X**	**-**	**X**	**-**	**-**
Wu et al. **[18]**	**L**	**H**	**L**	**H**	**H**	**H**	**M**	**L**	**-**	**X**	**L**	**L**	**H**	**H**	**-**	**X**	**-**	**-**	**X**
Ahmed et al. **[19]**	**L**	**M**	**L**	**M**	**M**	**L**	**L**	**M**	**-**	**X**	**M**	**H**	**H**	**H**	**-**	**-**	**X**	**-**	**-**
Lee et al. **[20]**	**M**	**M**	**M**	**H**	**H**	**M**	**H**	**M**	**X**	**-**	**M**	**L**	**L**	**H**	**-**	**X**	**X**	**X**	**-**
Wan et al. **[22]**	**M**	**M**	**H**	**M**	**H**	**M**	**H**	**H**	**X**	**-**	**H**	**L**	**M**	**M**	**-**	**-**	**-**	**X**	**-**
Khan et al. **[24]**	**L**	**M**	**M**	**H**	**H**	**L**	**M**	**M**	**X**	**-**	**M**	**L**	**M**	**M**	**-**	**-**	**-**	**X**	**-**
Proposed Model	**M**	**H**	**H**	**H**	**M**	**M**	**H**	**H**	**X**	**-**	**H**	**M**	**M**	**H**	**X**	**X**	**-**	**X**	**X**

Table 15.4: Comparison of related research work’s challenges, issues and proposals

NC: Needs of Cloud Computer users	**RR:** Resources management	**FM:** Federated methods
BPT: Big Data processing techniques	**VC:** Virtual Cloud	**RM:** Reinforcement methods
EES: Energy efficient scenario	**RC:** Real Cloud	**P:** Propose platform
AIT: AI techniques	**IS:** Integration Scenarios	**A:** Propose architecture
AIA: AI applications	**IA:** IoT applications	**F:** Propose framework
AII: AI interfaces	**BA:** Big Data analytics	**H:** High **L:** Low
AIC: AI Cloud-based platform	**BM:** Big Data management	**M:** Medium

Table 15.5: Table 15.4 parameters explanation

15.4.6 Comparative Analysis

Regarding the research we have made on previous works, we could analyze our findings in Table 15.4.

Specifically, Table 15.4 presents the basic model characteristics of former related works, compared with our proposed model. In particular, we can observe that most of the works do not contribute to technologies such as Federated Learning and Reinforcement Learning. In addition to this most of them do not associated to IoT Applications and the data produced from them. Another conclusion that we can observe is that the most of the related works implement platforms. Few of them try to find out solutions for the open issues and the needs of Cloud users, and more energy efficient systems for the operations. On the other hand, most of the related works contributed with the broader sense of Artificial Intelligence regarding novel technics, applications, interfaces, and platforms. Consequently, regarding the findings listed in Table 15.4, we can observe that our proposed scenario try to collaborate with novel scenarios such as Federated and Reinforcement Learning systems, and in addition embeds technics that implements a more energy efficient system that aims to offer more useful and efficient management of industrial IoT Big Data, combining both Cloud Computing, Big Data and Internet of Things in a novel framework. Table 15.5 presents the meanings of the abbreviations listed in Table 15.4.

15.5 Chapter Summary

This paper surveyed the multiple open challenges and issues in the field of energy-efficient industrial IoT-based Big Data management in Cloud environments, and in particular aspects and challenges arises from the fields of Artificial Intelligence scenarios of the Cloud infrastructures, Artificial Intelligence techniques of industrial IoT-based Big Data Analytics in the Cloud environments and Federated Learning Cloud systems try to be clarified. Take into account that Reinforcement Learning is a novel technique that allows large data centers such as Cloud data centers to affect a more energy-efficient resource allocation, we proposed an architecture that tries to combine the features offered by the several Cloud Providers in order to emerge and achieve an Energy-Efficient industrial IoT-based Big Data Management Framework (EEIBDM) established outside of every user, in Cloud environment. As a result, the major goal of this paper is the formulation of the various aspects of the resource allocation issue, considered from Reinforcement Learning scenario, while interacting with the Cloud environment with the aim to achieve an optimal decision. To achieve this, we proposed an algorithm for deliver the energy consumption of CPU through the evaluation of EEIBDM framework.

As a case study for the future, we are oriented to get involved to our proposed system framework the aspect of security, in order to achieve both energy-efficient ad secure environment in the Cloud for managing industrial IoT-based Big Data, with the help of novel learning techniques of Reinforcement and Federated Learning. Thus, this proposed framework could be used in places such as hospitals, universities and repositories of legal cases, in order to have a more secure environment, in addition to the most energy efficient environment.

15.6 Chapter References

[7] G. S. Aujla, N. Kumar, “MEnSuS: An efficient scheme for energy management with sustainability of cloud data centers in edge-cloud environment”, Future Generation Computer Systems, vol. 86, pp. 1279-1300, September 2018.

[8] A. Al-Dulaimy, W. Itani, R. Zantout, A. Zekri, “Type-aware virtual machine management for energy efficient cloud data centers”, Sustainable Computing: Informatics and Systems, vol. 19, pp. 185-203, September 2018.

[9] Md A. Khan, A. Paplinski, A. M. Khan, M. Murshed, R. Buyya, “Dynamic Virtual Machine Consolidation Algorithms for Energy-Efficient Cloud Resource Management: A Review”, Springer, Sustainable Cloud and Energy Services, chapter 6, pp. 135-165, September 2017.

[10] I. Weber, H. Wada, A. Fekete, A. Liu, L. Bass, “Automatic Undo for Cloud Management via AI Planning”, in Proceedings of the Eighth USENIX conference on Hot Topics in System Dependability (HotDep'12), October 2012.

[11] L. E. Brown, D. Kauchak, “Educational Advances in Artificial Intelligence”, AI Magazine, vol. 34, no. 4, p. 127, September 2013. [DOI: 10.1609/aimag.v34i4.2508]

[12] P. Rad, M. Roopaei, N. Beebe, M. Shadaram, Y. A. Au, “AI Thinking for Cloud Education Platform with Personalized Learning”, In Proceedings of the 51st Hawaii international conference on system sciences (HICSS 2018), 3-6 January 2018, Hawaii, USA. [DOI: 10.24251/HICSS.2018.003]

[13] C. Wu, R. Buyya, K. Ramamohanarao, “Big Data Analytics = Machine Learning + Cloud Computing”, Chapter 1 in “Big Data: Principles and Paradigms”, R. Buyya, R. Calheiros, and A. Dastjerdi (eds), Morgan Kaufmann, Burlington, Massachusetts, USA, 2016, arXiv:1601.03115.

[14] E. Ahmed, I. Yaqoob, I. A. T. Hashem, I. Khan, A. I. A. Ahmed, M. Imran, A. V. Vasilakos, “The role of big data analytics in Internet of Things”, Elsevier, Computer Networks, vol. 129, Part 2, pp. 459-471, December 2017. [DOI: 10.1016/j.comnet.2017.06.013]

[15] J. Lee, H. Davari, J. Singh, V. Pandhare, “Industrial Artificial Intelligence for industry 4.0-based manufacturing systems”, Elsevier, Manufacturing Letters, vol. 18, pp. 20-23, October 2018. [DOI: 10.1016/j.mfglet.2018.09.002]

[17] J. Wan, J. Yang, Z. Wang, Q. Hua, “Artificial Intelligence for Cloud-assisted Smart Factory”, IEEE Access, vol. 6, pp. 55419-55430, September 2018. [DOI: 10.1109/ACCESS.2018.2871724]

[18] S. Khan, K. A. Shakil, M. Alam, "Cloud-Based Big Data Analytics - A Survey of Current Research and Future Directions", Springer, Big Data Analytics, pp. 595-604, October 2017. [DOI: 10.1007/978-981-10-6620-7_57]

[19] H. Manjunatha, E. T. Esfahani, "Application of Reinforcement and Deep Learning Techniques in Brain–Machine Interfaces", Springer, Advances in Motor Neuroprostheses, pp. 1-14, April 2020.

[20] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, D. Hassabis, "Reinforcement Learning, Fast and Slow", Elsevier, Trends in Cognitive Sciences, vol. 23, issue 5, pp. 408-422, May 2019.

[21] M. Alhamad, T. Dillon, E. Chang, ''Conceptual SLA Framework for Cloud Computing". In Proceedings of the 2010 4th IEEE International Conference on Digital Ecosystems and Technologies (DEST), pp. 606 -610, 13-16 April 2010, Dubai, United Arab Emirates.

[22] G. Koutitas, P. Demestichas, "Challenges for energy efficiency in local and regional data centers", Journal on Green Engineering, pp. 1.32, October 2010.

Chapter 16

Conclusions & Future Directions

In this chapter presented the main objectives and outcomes of my research. As the last chapter of my dissertation summarizes all the most important findings of each work presented in previous chapters, as well as what conclusions I came to. In a large part of these researches presented in my dissertation I had significant help from my supervising professor, Dr. Kostas E. Psannis, but also from other members of the research team Mobility2Net with whom we had a very good collaboration in research works that were made in combination. The conclusions section presents all the important conclusions that were drawn from the work I did to complete my dissertation. Whereas, Future Directions section presents all the open research questions that could be studied and explored in future work and draw additional useful conclusions.

16.1 Conclusions

All the major and important conclusions will be presented in this section based on the order in which the respective works were presented in the previous sections.

Initially, I surveyed BD and CC technology and their basic characteristics, with a focus on the security and privacy issues of both technologies. Through this study, I have tried to combine the functionality of the two aforementioned technologies to examine the frequent characteristics and to discover the benefits related to the security issues of their integration. Count on this, as the first main goal of my research was to find novel ways to achieve better integration of BD and CC, with a focus on security algorithms and all the challenges that the two aforementioned technologies faced on the security level. However, regarding the rapid development of both technologies, the security issue must be solved or reduced to a minimum to have a better integration model. Thus, all the security challenges that were surveyed in this proper work were the sector for further research as a case study, to minimize them.

Subsequently, I surveyed IoT technology, with an explanation of its operation and use, and the main features of MCC and its trade-offs. Additionally to these, I have studied Big Data Applications and some of their basic features, along with the contribution of the IoT technology, and the MCC technology to Big Data Applications. The findings of this study and the exploration of the contribution provided by the Internet of Things features, and by MCC features in dealing with the basic characteristics of the Big Data Applications, are shown in tables 3.1 and 3.2, respectively. So, based on the findings of tables 3.1 and 3.2 can be inferred that IoT and MCC possess features that could be beneficial for the use of Big Data Applications.

In the next research that has been made, I further studied IoT, with an explanation of its operation and use, and also the main features of CC and its trade-offs. Since CC refers to an infrastructure where both data storage and data processing happen outside of the mobile device, and the IoT is a new technology that is growing rapidly in the field of telecommunications, and especially in the modern field of wireless telecommunications, I reached the next step of my research. The main goal at this stage is the interaction and cooperation between things and objects sent through the wireless networks is to fulfill the objective set to them as a combined entity. Based on the technology of wireless networks, both the technologies of CC and IoT develop rapidly, so I have combined CC and IoT to examine the common features and to discover the benefits of their integration. Moreover, the contribution of CC to the technology IoT, shows how the CC technology improves the function of the IoT. Finally, the security challenges of the integration of IoT and CC were surveyed through the proposed algorithm model presented in the work titled "*Secure integration of IoT and Cloud Computing*", and also there is a presentation of how the two encryption algorithms which were used, contribute in the integration of IoT and CC. Based on these findings, it could be done further research to reduce to minimum security issues of both CC and IoT to have a better integration model.

Additionally, in another work made at the same time, I surveyed BD and CC with a focus on the security and management issues of both technologies, and so I have tried to combine CC and BD to examine the related features and to discover the benefits of their integration. Thus, through the research titled "*Secure integration of Cloud Computing and Big Data*" I have found how the CC technology improves the function of BD, and also I surveyed the security challenges of the integration of BD and CC and proposed a novel security model, count on the previous encryption algorithms AES, RC5, RSA.

Another technology studied in my research is MCC, which refers to an infrastructure where data, applications, and information could be processed through a mobile device, but simultaneously outside of the mobile device. The main objective of the use of MCC is to decrease the use of stronger hardware and to have the access to data and applications, and many times to more computational power, from every place and at any time, through a mobile device. Due to this and regarding the huge benefits of MCC, I have tried to achieve a more safe and trusted environment for the MCC users to operate the functions, and transfer, edit and manage data and applications. I realized that this could be achieved by proving a novel method count on the AES encryption algorithm, which is the most relevant encryption algorithm to a Cloud environment. Therefore, I have tried to define the most important issues and challenges in the field of MCC technology by presenting a number of the most significant works related to MCC through the last recent years. This research could lead to the solution or to the reduction of minimizing all the challenges and issues that MCC and the related to its technologies faced on the security level, to have a better and safer model.

Following previous research about CC and IoT, the main objective of the interaction and cooperation between things and objects sent through the wireless networks is to fulfill the objective set to them as a combined entity, to achieve a better environment for the use of BD. As a result, in the work titled "*Secure Machine Learning scenario from Big Data in Cloud Computing via Internet of Things network*" a combination of CC and IoT achieved, to examine the familiar characteristics, and to discover the benefits of their integration to secure the use and the transmission of BD. Thus, the security challenges of the integration of IoT and CC were surveyed through the proposed algorithm model of this work, and additionally, there is a presentation of how the two encryption algorithms which were used, contribute to the integration of IoT and CC as base technologies for BD.

Also, following another previous research, I have tried to combine the functionality of BD and CC to examine the frequent characteristics, and to discover the benefits related to security issues of their integration for once more, to present a new method of an algorithm that can be used to improve CC's security through the use of algorithms that can provide more privacy in the data related to BD. Count on this further researches on the integration of those two technologies can be done then, and also studies of having a huge improvement of security and privacy issues of CC and BD to have better use of them.

Regarding the previous findings of my research, I have studied out the need of discovering new methods of technological support in many sciences by the researchers of CC. As part of these researches, in the work "*Algorithms for efficient digital media transmission over IoT and cloud networking*", to achieve a type of network that will provide more intelligent media-data transfer, I have studied, in collaboration with other researchers of my research team, new technologies, and the use of various open-source tools, such as CC analyzers and simulators. Tools like these are useful for studying the collection, storage, management, processing, and analysis of large volumes of data. Furthermore, the simulation platform used in this research is CloudSim and operates in the Eclipse environment. Additionally, after measuring the network performance with CloudSim, in the context of this work, we used the Cooja emulator of the Contiki OS to confirm and access more metrics and options. As a result, we implemented a network topology from a small section of the script of CloudSim with Cooja, so that we can simulate a single network segment. The results of the experiment show that there are not duplicated packets received through the network. So, this leads the whole research team to the conclusion that further examination of the simulation analysis of the network performance in CloudSim simulator needed, and other simulation platforms, to have a better and improved contribution of IoT with the additional "*help*" of the CC for better transmission of high-quality data, could be a start point for better and more efficient media data transmission.

Moreover, through the work "*Security, Privacy & Efficiency of Sustainable Cloud Computing for Big Data & IoT*" I have tried to establish an architecture relying on the security of the network to improve the security and privacy issues of CC, IoT, and BD. The simulations which have been made in this work present the proposed solution which is the installation of a security "*wall*" between the Cloud Server and the different users on the Internet. This proposal aims to eliminate the privacy and security issues that need to be faced, and considered that CC provides efficiency in privacy issues of the network, where bits are transferred through time. Also, the security challenges of the integration of IoT and CC surveyed intending to provide an architecture relying on the security of the network to improve the security issues and count on this state that CC could offer a more "*green*" and efficient fog environment for sustainable computing scenarios. That statement leads me to more substantial and constructive researches then.

In another significant work, I have tried to study and made my research related to the novel 6G communication networks. Thus, in "*IoT-based Big Data secure management in the Fog over a 6G Wireless Network*" I aimed to offer a better scenario on a Cache Decision System (CDS) of a Smart Building, established on a University campus as a case scenario, through a wireless network. Through this wireless network, I can combine the functions of IoT, CC, EC, and BD which play a vital role in the telecommunications field. The wireless network of the intelligent-smart building was based on 6G technology, with all the benefits this technology offers. Thus, the main purpose of this work was to propose a novel CDS through a 6G wireless network, which will offer users, safer and efficient environment for browsing the internet, sharing, and managing large-scale data (BD) in the fog. The proposed CDS consisted of two types of servers, one Cloud Server and one Edge Server. To come up with this proposal, I have studied related caching system scenarios from former works, which are listed and presented in more detail in this work.

These previous works inspired me to deal with another BD perspective. The streaming high-quality video can be represented as BD and could be managed, be processed, and be transferred through Cloud environments. So, I have surveyed the three aforementioned technologies to find the common features of their use and to propose an operation that would help the issue of streaming high-quality video. Based on these results, I have tried to propose an efficient algorithm for advanced scalable Media-based Smart Big Data (3D, Ultra HD) on Intelligent Cloud Computing systems. Based on the performance evaluations that have been made can be demonstrated that the proposed encoding algorithm outperforms the traditional HEVC standard. Thus, by adopting this proposed method could be assumed that it can be used and integrated into HEVC without violating the standard.

Except for the security and privacy issues that have been tried to be solved through various and novel proposed methods, another field that has been carried out and studied in my research is the energy-efficient issues of CC. After I came through the previous researches, the CloudSim simulator as the ideal Cloud simulator for my research, I investigated the related works on the field of Green Cloud Communications where I have tried to offer a novel energy-efficient and green Cloud environment for data management and processing through the CloudSim simulator. Thus, count on the basic aspects of CloudSim, I proposed a system framework for better use of BD management, based on the idea of a Cloud federated network. I tried to offer an algorithm approach to this scenario, by proposing a novel model for achieving an energy-efficient resource allocation technique for BD management in the Green Cloud environment. Through the experimental results of this work shown that the proposed model has immense potential as it offers significant performance gains regarding cost-saving and better data management under large workload scenarios.

Through this previous research, the idea of using the federated scenario along with CC came up. Subsequently, through the work titled "*InFeMo: Flexible Big Data management through a federated Cloud system*" I have presented and described a new system architecture based on CC, and count on the novel scenario of Federated Learning, which called Integrated Federated Model - InFeMo. This model incorporates all Cloud models with a federated learning scenario, as well as other technologies that may have integrated use with each other. The major motivation of InFeMo is to provide more efficient system architecture and environment for the users' data management. The efficiency of this proposal counts on its operation, because it decreases the number of rounds of communication that are needed to train a scenario model by using a federated Cloud system, and as a result, it makes the user that uses this system wait for less. The system's federated algorithm relies on the advantages of the former models of FedAvg and CO-OP algorithms. Consequently, I ended up with my new scenario that merges these two algorithms aiming to have a more efficient model, which selects the training model depending on each occasion. This work tried to fill the scientific gap in the field of federated cloud systems and can lead to more researches and experiments to achieve and explore the new opportunities arising in this field. Thus, based on this research made and the comparative analysis, no other architecture model has been studied and proposed so far, which incorporates all Cloud models with a federated scenario, as well as other technologies that may have integrated use with each other. So, I was motivated to work in this field trying to find out new aspects that could lead to efficient and more secure communication between the user and the central server. Also, I have tried to involve an IoT scenario of sensors and a SB scenario to find out new methods and aspects that arise here.

Finally, another scenario that involves Federated Learning with CC and Reinforcement Learning carried out. The major goal of the last steps of my research was to find a more secure and "*green*" Cloud environment. Through the work titled "*EEIBDM: A Reinforcement & Federated Learning scenario for Efficient Industrial IoT-based Big Data Management in Cloud*" surveyed the multiple open challenges and issues in the field of energy-efficient industrial IoT-based BD management in Cloud environments, and particular aspects and challenges arises from the fields of AI scenarios of the Cloud infrastructures, Artificial Intelligence techniques of industrial IoT-based BD Analytics in the Cloud environments and Federated Learning Cloud systems try to be clarified. In addition to this and taking into account that Reinforcement Learning is a novel technique that allows large data centers such as Cloud data centers to affect a more energy-efficient resource allocation, I proposed an architecture that tries to combine the features offered by the several Cloud Providers to emerge and achieve an Energy-Efficient IoT-based Big Data Management Framework (EEIBDM) established outside of every user, in Cloud environment. Based on this, the major goal of this work, count on the major's dissertation work, was the formulation of the various aspects of the resource allocation issue, considered from the Reinforcement Learning scenario, while interacting with the Cloud environment to achieve an optimal decision. To achieve this, I have proposed an algorithm for delivering the energy consumption of CPU through the evaluation of the EEIBDM framework.

16.2. Future Directions

Certainly, it has been more aspects to be studied in the field of BD management, security, processing, and transmission through CC that I would like to further investigate in the future. Also, more research aspects of how Cloud could lead us to a more "*green*" and efficient scenario need to be done. Thus, I will be able to make further research in these fields to continue the former work I have already made.

I hope through the next steps and as a continuation of my research to make more simulations to have better accuracy in my experimental results, and more data transfer scenarios have to be made through the simulators providing results counting not only in data transmission but also in network efficiency and support. Also, considering that CC is a novel technology that is constantly evolving and becoming more and more necessary, in the sector of communications and not only, more study needs to be made about its operation and how CC interacts and integrates in a better way with other technologies such as IoT and BD. So, this can be one of the fields of future research.

Moreover, it already is planned, as another case study, to improve my proposed system and investigate for an even better Case Decision Algorithm for the Cache Decision System in which all necessary configurations will be taken into consideration, such as every service that will be provided by the Intelligent Building to the users that have access on it. Consequently, this research could be a start point for a better and more efficient wireless networking scenario, for managing and sharing BD on a SB. Furthermore, as an extension of the proposed scenario, it is planned to be tested on a Smart Hospital, due to the meaning of the data used and transferred on a hospital, which are needed to be secured and immediately accessible. Thus, as a case study for the future, I can implement, along with the other members of the research team Mobility2Net, an Extremely Scale Analytics System (ESAS), relying on the experimental results of this dissertation. The ESAS would be installed in the central server of the proposed CDS, and so it will take advantage of the system’s efficient operation settled in the SB.

Regarding energy efficiency, it is oriented to involve to the proposed framework the major issue of security, to provide a Green and Secure Cloud environment, through the scenarios of federated learning and reinforcement learning. So, as a case study for the future, is to get involved in the proposed EEIBDM system framework the aspect of security, to achieve both an energy-efficient and secure environment in the Cloud for managing industrial IoT-based BD, with the help of novel learning techniques of Reinforcement and Federated Learning. Based on this, the EEIBDM framework could be used in places such as hospitals, universities, and repositories of legal cases, to have a more secure environment, in addition to the most energy-efficient environment.

Finally, there are also other areas where my proposed InFeMo model could be applied, beyond the academic community, offering multiple benefits. As already mentioned in the specific chapter, the academic community can offer users a more efficient environment, offering less waiting time and ease in the mass management of their data. Regarding the health sector, where the proposed InFeMo model could also be applied, as it would mainly facilitate the medical staff in the easier access to the sensitive medical data and their analysis, which can be performed in the specific Cloud environment of this system, more immediate and efficient. Thus, for example, the doctor will be able to receive the data needed through the smartphone directly, and also due to the use of the federated scenario to enable data analysis applications to provide data immediately and quickly. Additionally, another area of application that the proposed model could be involved in the industry. There it would facilitate the most efficient supervision of the production process of a production chain. Based on the federated scenario of the proposed model, the managers of each related part of the production will be able to receive data analysis components, but also to have faster and more direct access to them. Moreover, even in the administrative part of an industry, it could help in better and faster access to data, by giving the users the ease of using a Cloud infrastructure without the necessary choice of a specific provider. These could be the next steps of the future continuation of my current research.

Appendix 1

In Appendix are listed all the peer-reviewed papers related to the main scope of the dissertation but not part of basic research of it.

Papers are listed in chronological order of publication.

1. Work Title: Architecture for Security in IoT Environments

C. Stergiou, K. E. Psannis, A. P. Plageras, G. Kokkonis, Y. Ishibashi, "Architecture for Security in IoT Environments", in Proceedings of 26th IEEE International Symposium on Industrial Electronics, 19-21 June 2017, Edinburgh, Scotland, UK. [DOI: 10.1109/ISIE.2017.8001447]

Abstract

The focus of this paper is to propose an integration between Internet of Things (IoT) and Video Surveillance, with the aim to satisfy the requirements of the future needs of Video Surveillance, and to accomplish a better use. IoT is a new technology in the sector of telecommunications. It is a network that contains physical objects, items, and devices, which are embedded with sensors and software, thus enabling the objects, and allowing for their data exchange. Video Surveillance systems collect and exchange the data which has been recorded by sensors and cameras and send it through the network. This paper proposes an innovative topology paradigm which could offer a better use of IoT technology in Video Surveillance systems. Furthermore, the contribution of these technologies provided by Internet of Things features in dealing with the basic types of Video Surveillance technology with the aim to improve their use and to have a better transmission of video data through the network. Additionally, there is a comparison between our proposed topology and relevant proposed topologies focusing on the security issue.

Keywords—Internet of Things, video surveillance, IoT, monitoring, network topology, architecture.

I. Introduction

A number of modern mobile devices, like mobile phones, PDAs, laptops and others, become ubiquitous in recent years and people into the era of pervasive computing [1]. All these devices could be used with the aim to find out useful information when we are on the road and when we are travelling. This procedure can help us to define monitoring. Thus, "Monitoring is the act of listening, carrying out surveillance on, and/or recording the emissions of one's own or allied forces for the purpose of maintaining and improving procedural standards and security, or for reference, as applicable" [2].

Regarding this definition it is proved that monitoring related to surveillance. So, also, we could define surveillance, as a related part of technology in this work. Surveillance is "the close observation of the behaviour, the activities, or other changing information" [3] [4]. Sensors and cameras or other compatible devices are necessary for the surveillance with the aim to do the monitoring. With the use of this technology observation at a distance is possible, using electronic equipment [4] or stealing electronically transmitted information which may include simple, relevant technology methods.

Furthermore, in telecommunication fields there is a new technology called Internet of Things (IoT) [5]. The next major step in the recent technology field is the IoT technology, but however with the major difference that brings enormous changes in business functionality [6] [7]. In order to fully exploit these two technologies, it is mandatory to combine them so as to achieve the optimisation of surveillance technology through the use of the Internet of Things technology [8] [9].

The rest of the paper is organised as follows. In section 2 there is a review of the related research which deals with the monitoring urban areas throw modern networks. Section 3 presents and illustrates the proposal of a contribution of the Internet of Things technology in the function of Video Surveillance with the aim to offer a new topology paradigm. In Section 4 there is a comparison between our proposed topology and other related proposed architectures-topologies. Finally, section 5 provides the conclusions of the current paper and offers new possibilities for the development of future work.

II. Related Review

For the purpose of this paper we study and analyse previous studies in monitoring urban areas throw modern networks and we examine existing work proposed both in the literature and on the Internet. Below presented the papers we have studied with their main objective.

There are various works for the monitoring urban areas throw modern networks. A large number of several works related to monitoring urban areas throw modern networks the last two years. To begin with, the authors of [10] introduces the Shadow Security Unit, a low-cost device deployed in parallel with a PLC or Remote Terminal Unit (RTU), being able to transparently intercepting its communications control channels and physical process I/O lines with the aim to continuously assess its security and operational status. The device that proposed in [10], regarding the existing control network, does not require considerable changes, in order to be capable of work in standalone or integrated within an ICS protection framework. Also, by the work that has been made in [11], the authors propose an innovative approach for the development of software for modelling of decentralised intelligent systems for security monitoring and control in power systems. The novelty of [11] is to joint use the modern computing environments. Also, the proposed intelligent system was tested on the modified 53-bus IEEE power system. The main aim of [12] is to describe an innovative security system able to localise and classify audio sources in an outdoor environment. The primary intended use of the proposed security system is for security monitoring in serve scenarios, and it has been designed to cope with a large set of heterogeneous objects, including weapons, human speakers, and vehicles. Also, in [12] after the presentation of the details of the system's design, with a particular emphasis on the innovative aspects that are introduced with respect to the state-of-the-art, the authors offer an extensive set of simulations in order to show the effectiveness of the proposed architecture. At the end the authors conclude be describing the current limits of the system, and the projected further developments. The current knowledge in the regard of the use of different tools needed in order to monitoring atmospheric pollution extended in [13]. The chemical response of the lichen Ramalina celastri was evaluated through physiological parameters and sulfur accumulation in relation to the SO2 and NO2 concentrations present in the air at the monitoring sites with different emission sources, with the aim to assess the atmospheric pollution in urban environments. Regarding this, it was possible to create different levels of air quality using simultaneous measurements of gaseous pollutants in the air and of parameters for the exposed biomonitor, as well as to determine the relationship between them and their society with the different emission sources present. In addition, in [14] discussed that in regions with a mild climate, pesticides are often used around homes for pest control. Pesticide use in residential areas linked to aquatic toxicity in urban surface water ecosystems, and suggested dust particles on a paved surface as an important source of pesticides by the recent monitoring studies which have been made. With the aim to be tested the hypothesis that dust on hard surfaces is a significant source of pesticides; the authors of [14] evaluated spatial and temporal patterns of current-use insecticides in Southern California, and further explored their distribution as a function of particle sizes.

The [15] reports on the first results of a long-term UFP monitoring network, set up in Amsterdam (NL), Antwerp (BE), Leicester (UK) and London (UK), with the aim to gain a better understanding on the spatiotemporal alteration of ultrafine particles (UFPs) in urban environments. Furthermore, the authors of [15] in order to represent the extreme rainfall-runoff events, the deterministic distributed hydrological modelling is gaining interest both with the increase of the computation facilities and the availability of data especially the topography inputs. Also, in [16] the simulation results of four deterministic hydrological models with different topography resolution (300m, 150m, 75m) for the Var basin, France (2800km2) are analysed with the aim to evaluate the influences on the simulation accuracy. The results of sensitivity analysis indicate the threshold value of the topography resolution on the model simulation with the consideration of both the sufficient accuracy and the reasonable simulation time to cover the extreme rainfall-runoff event in 1994. In [17] the authors introduce a framework for precise vehicle localisation in dense urban environments that are characterised by high rates of dynamic and semi-static objects. The proposed localisation method of [17] is particularly designed for handling the inconsistencies between map material and sensor measurements. The evaluation results of this work show the superior performance of the proposed approach compared to another state-of-the-art localisation algorithm for a challenging urban dataset.

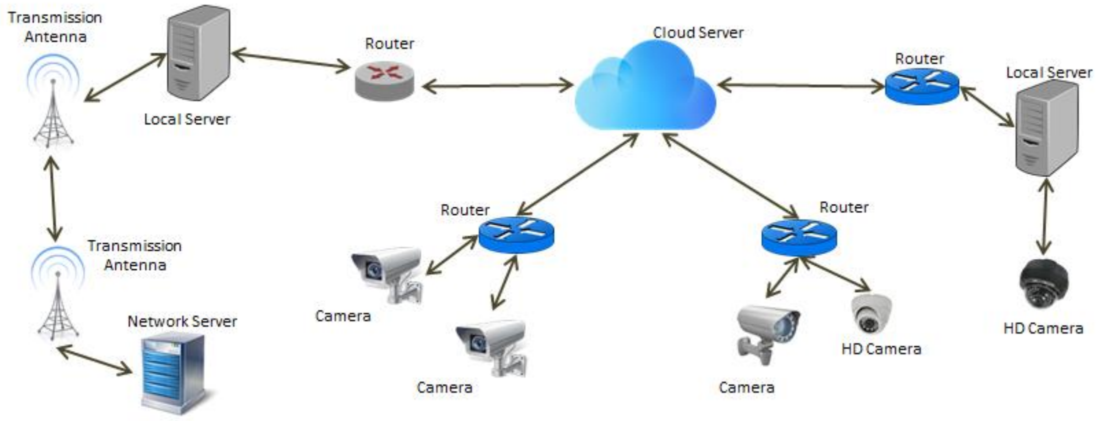


Figure 1: Proposed topology paradigm.

Video Surveillance	Smart solution in the bucket of transport	Smart power grids incorporating more renewable	Remote monitoring of patients	Sensors in homes and airports	Engine monitoring sensors that detect & predict maintenance issues
Computer		X	X		X
Telephones			X		X
Cameras		X	X	X	X
Biometric		X	X		
Data mining and profiling	X		X		X
RFID and geolocation devices	X		X	X	X

Table 1: Contributions of Internet of Things in Video Surveillance.

III. Topology Proposal

Concerning our research of the Related Research Review section, we developed the following conclusions as a proposal of IoT's contribution in Video Surveillance. A major issue of the Video Surveillance technology is the transmission of data through the video recorder devices and how those devices should be set up with the aim to have a better use of remote control.

As a solution to this problem, we propose an innovative topology paradigm that combines the advantages of the use of IoT and the characteristics of Surveillance. This proposed topology is a hybrid topology of ring and star topologies. In this topology we could succeed a reliable network in error detecting and troubleshooting, we could scalable the size of the network as in can be increased easily, and additionally, this topology offers flexibility and provides a more effective network.

Furthermore, as a combination of two topologies we can operate this network both as a star-topology-network so as a ring-topology-network, as well as a separate type of networks. By using routers with the aim to have single management network sectors, we can achieve a different type of topology use. Figure 1 presents a paradigm of the proposed topology using two types of video surveillance cameras (simple quality surveillance camera and HD quality surveillance camera). The data transmitted from the cameras to the Cloud Server with the useful help of IoT technology and from the Cloud Server transmitted to another local server, and finally we can have all the transmitted data in the storage system of the network server. As it is shown each router could be able to serve a huge number of network cameras, connected to each other with different ways. Also, in this type of network topology, Local Servers can be used inside the small networks as administrators of network cameras. Cloud Server could provide primarily the important role of the storage system, and afterwards could act as data manager that receives these data with the aim to transmit them to the Network Server. Between the Cloud Server and the Network Server, also could interpolate another Local Server in order to clarify and transmit the data to its final destination, which is the Network Server.

An important improvement in the operation of this topology is analysed and described by the following equation:

$$DS = (TD + VD) - PL \quad (1)$$

Equation (1) demonstrates Data send (DS) through the network. This data results by the product of the quantity of the Transmitted data (TD) and the quantity of Video data (VD) deducting the quantity of the Packet Loss (PL). By this equation and regarding the number of nodes existing in the network, we can produce the total amount of data which transmitted through the network. Thus, this calculated by (2):

$$TDS = DS_1 + DS_2 + ... + DS_n \quad (2)$$

Moreover, through our research, we detect that another major issue of the Internet of Things and Video Surveillance technology is the event detection problem in noisy environments for a multimedia monitoring application which is solved with the detection of the abnormality in continuous audio recordings of public places [18]. Regarding the combination of the aforementioned technologies, Table 1 lists the characteristics of the technology of Things, with regard to the convenience it provides. It also demonstrates some of the types of Video Surveillance technology which relates more, in our opinion, to the Internet of Things. Table 1 has the purpose to show which of the specific characteristics of the IoT technology pertain to, and improve the particular types of the Video Surveillance. As we can observe, Cameras and RFID devices are the Video Surveillance types which are affected more by the characteristics of the IoT technology. In contrast, the Biometric is the type of Video Surveillance influenced less by the characteristics of IoT technology.

Video Surveillance Architectures - Topologies	Efficiency	Security	Easy Installation	Transmission Speed	Quality of Communication	Data Privacy
Hierarchical Video Surveillance Architecture [20]			X	X	X	
Wireless Mesh Networks [21]	X	X		X		
Scalable & Robust Framework [22]	X			X	X	
Topology estimation for thousand-camera surveillance networks [23]	X		X	X	X	
Microphone array based classification [12]		X			X	X
IoT-based Surveillance System [24]	X	X	X		X	
Proposed Topology		X	X		X	X

Table 2: Architectures-Topologies Comparison.

Additionally, the third technology that takes place in the IoT's contribution in Video Surveillance is Cloud Computing. Cloud Computing is a technology that could be set as a base technology in the use of both IoT along with Video Surveillance. Cloud Computing is the outcome of interdisciplinary approaches, combining mobile computing and Cloud Computing [19]. Since the Cloud Computing is used as a basis for both IoT and Video Surveillance, we can claim that the former improves the function of the IoT, and influences the different types of Video Surveillance technology.

IV. Architecture Comparison

The study of previous works cites us relevant architecture and topology proposals for a Video Surveillance network, which on several occasions supported and combined with other technologies, such as Internet of Things. In this section we will make a comparative study of the proposal made in this work and proposals made by other relevant works.

On the study conducted we singled out six previous architecture-topology proposals relating to Video Surveillance technology. Here, there will be a comparison between the features and the benefits of each proposal. As we can observe from Table 2 most former works deal with the Quality of Communication, and as the second characteristic that deal with is Security. Thus, the main purpose of these works is to provide secure and quality communication architecture. Comparing our proposed topology to the others we can realise that it contributes more security and privacy issues. It has certainly a disadvantage in relation to the others, as regards the Transmission Speed and the Efficiency.

In addition, the proposed topology of this work could mainly be applied in big buildings, in which there are installed systems of surveillance cameras. Buildings such this could also be defined as Smart Buildings instead of the specialised use of the surveillance system in conjunction with the IoT technology. Thus, the proposed topology can be described as an ideal topology for surveillance systems after using a combination of IoT and Cloud Computing technologies.

V. Conclusion

With regard to the use of the Video Surveillance and the future needs of this technology, there has been a combination of Video Surveillance technology with Internet of Things technology in order to take advantage of the IoT benefits and improve the use of Video Surveillance. The discussion of this contribution proposes an innovative topology paradigm which could offer a better use of IoT technology in Video Surveillance systems. Also, the contribution of these technologies provided by Internet of Things features in dealing with the basic types of Video Surveillance technology is summarised in Table 1. Additionally, there is a comparison between our proposed topology and relevant proposed topologies focusing on the security issue. Finally, as a future research, we suggest a further examination of the types of Video Surveillance which could be improved from the contribution of the technology of Internet of Things with the additional ‘help’ of the Cloud Computing technology.

VI. References

[1] Uichin Lee et al, «MobEyes: Smart mobs for urban monitoring with a vehicular sensor network,» *IEEE Wireless Communications,* pp. 1-15, 1/11/2006.

[2] Dictionary.com, "Dictionary.com," Dictionary.com, 1/1/2012. [Online]. Available: http://www.dictionary.com/browse/monitoring. [Accessed 2/12/2016].

[3] M. Maximino et al, "Journalist's Resourse, Research on today's news topic," 11/2/2014. [Online]. Available: http://journalistsresource.org/studies/government/criminal-justice/surveillance-cameras-and-crime. [Accessed 6/3/2016]

[4] J. M. Batalla et al, “Adaptive Video Streaming: Rate and Buffer on the Track of Minimum Rebuffering”, IEEE Journal on Selected Areas in Communications, vol. 34, Issue 8, pp. 2154-2167, 1/8/2016.

[5] Christos Stergiou, Kostas E. Psannis, Byung-Gyu Kim, Brij Gupta, Secure integration of IoT and Cloud Computing, Elsevier, Future Generation Computer Systems, December 2016 (http://www.sciencedirect.com/science/article/pii/S0167739X1630694X).

[6] Sandip Roy et al, "A Fog-Based DSS Model for Driving Rule Violation Monitoring Framework on the Internet of Things," International Journal of Advanced Science and Technology, pp. 23-32, 1/3/2015.

[7] Jordi Mongay Batalla and Piotr Krawiec, "Conception of ID layer performance at the network level for Internet of Things," *Pers Ubiquit Comput,* no. 18, pp. 465–480, 28/4/2013.

[8] George Kokkonis, Kostas E. Psannis, Manos Roumeliotis, and Yutaka Ishibashi, Efficient algorithm for transferring a real-time HEVC stream with haptic data through the internet, Journal of Real-Time Image Processing, May 2015. http://link.springer.com/article/10.1007/s11554-015-0505-7.

[9] George Kokkonis, Kostas E. Psannis, Manos Roumeliotis and Dan Schonfeld, Real-time wireless multisensory smart surveillance with 3D-HEVC streams for internet-of-things (IoT), Jounral of Supercompuitng, 2016, (http://link.springer.com/article/10.1007%2Fs11227-016-1769-9)

[10] Tiago Cruz et al, «Improving Network Security Monitoring for industrial control systems,» σε *Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on*, Coimbra, Portugal, 2015.

[11] Daniil Panasetsky et al, "Development of software for modelling decentralized intelligent systems for security monitoring and control in power systems," *PowerTech, 2015 IEEE Eindhoven*, pp. 1-6, 29/6/2015.

[12] Simone Scardapane et al, "Microphone array based classification for security monitoring in unstructured environments," *International Journal of Electronics and Communications (AEÜ),* no. 69, pp. 1715-1723, 11/11/2015.

[13] A.C. Mateos & C.M. González, "Physiological response and sulfur accumulation in the biomonitor Ramalina celastri in relation to the concentrations of SO2 and NO2 in urban environments," *Microchemical Journal,* no. 126, p. 116–123, 1/3/2016.

[14] Jaben Richards et al, "Distribution of pesticides in dust particles in urban environments," *Environmental Pollution,* no. 214, p. 290298, 7/4/2016.

[15] J. Hofman et al, "Ultrafine particles in four European urban environments: Results from a new continuous long-term monitoring network," *Atmospheric Environment,* no. 136, pp. 68-81, 8/4/2016.

[16] Qiang MA et al, "Assessment of High Resolution Topography Impacts on Deterministic Distributed Hydrological Model in Extreme Rainfallrunoff Simulation," in *12th International Conference on Hydroinformatics, HIC 2016*, Nice, France, 2016.

[17] Jan Rohde et al, "Precise vehicle localization in dense urban environments," in *19th International IEEE Conference on Intelligent Transportation Systems*, Rio de Janeiro, Brazil, 2016.

[18] C. Clavel, T. Ehrette & G. Richard, "Events Detection for an Audio-Based Surveillance System," Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on, pp. 1306-1309, 6/7/2005.

[19] Christos Stergiou & Kostas E. Psannis, "Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey," International Journal of Network Management, pp. 1-12, 11/3/2016.

[20] Sola O. Ajiboye et al, "Hierarchical Video Surveillance Architecture - A Chassis for Video Big Data Analytics and Exploration," in *Proceedings of SPIE - The International Society for Optical Engineering*, Falmer-Brighton, United Kingdom, 2015.

[21] F. Licandro & G. Schembra, "WirelessMesh Networks to Support Video Surveillance: Architecture, Protocol, and Implementation Issues," *EURASIP Journal onWireless Communications and Networking,* no. 2007, pp. 1-13, 30/1/2007.

[22] S. Dutt & A. Kalra, "A Scalable and Robust Framework for Intelligent Real-time Video Surveillance," Department of Electronics Engineering, Indian Institute Of Technology (BHU), Varanasi, India, 2016.

[23] Henry Detmold et al, "Topology Estimation for Thousand-Camera Surveillance Networks," IEEE, Adelaide, Australia, 2007.

[24] Andreas P. Plageras et al, "IoT-based Surveillance System for Ubiquitous Healthcare," in Industrial Electronics Society , IECON 2016 - 42nd Annual Conference of the IEEE, 22/12/2016.

2. Work Title: Security and Privacy of Big Data for Social Networking Services in Cloud

C. Stergiou, K. E. Psannis, A. P. Plageras, T. Xifilidis, B. B. Gupta, "Security and Privacy of Big Data for Social Networking Services in Cloud", in Proceedings of IEEE conference on Computer Communications (IEEE INFOCOM 2018), 15-20 April 2018, Honolulu, HI, USA.

Abstract

Big Data (BD) is of great importance especially in wireless telecommunications field. Social Networking (SNg) is one more fast-growing technology that allows users to build their profile and could be described as web applications. Both of them face privacy and security issues. In this paper, we survey SNg, BD and Cloud Computing (CC) technology and their basic characteristics, by concentrating on the security issues of those technologies. Specifically, we aim at combining the functionality of these two technologies (i.e Big Data and Social Networking) in a CC environment, so that we can analyze the common features and ascertain the advantages of their integration related to security issues. Through this research, we present a new system-framework-network in Cloud Environment through which users of various Social Networks (SNs) will be able to exchange data and information, and primarily large-scale data (Big Data). With our proposed system, we can achive greatly improve of the communication of SN users, and thus become more safe and accurate in a Cloud environment. More specifically, this system could be established as an intermediate communication node that could be utilized in order to improve the security of SNg's users through the use of algorithms that can provide more privacy in the data related to BD technology. Also, in this work we present some measurements and results relative to our proposed system use. Finally, the opportunity to create a database through which each user can view the statistics of his interaction with the SNg is further discussed.

Keywords— Cloud Computing, Big Data; Social Networking; Framework; System; Security; Privacy;

I. Introduction

SN is a structure consisting of sets of social, dyadic ties, and other social interactions between people. The SN perspective offers a set of methods for analyzing the structure of whole social entities as well as a variety of theories explaining the patterns observed in these structures [1]. SNs are "*self-organizing, emergent, and complex, such that a globally coherent pattern appears from the local interaction of the elements that comprise the system*" [2] (figure 1). Privacy concerns with SNg services is a subset of data privacy, involving the right of mandating personal privacy concerning storing, re-purposing, provide to third parties, and displaying of information connected with oneself through the Internet [3].

Figure 1. Social Networks Society.

CC constitutes a technology of internet services providing remote use of hardware and software. As a consequence, the users of CC could have access to information and data from any place at any time. In recent years, giant companies of the IT and software sectors investigate the services of CC. Furthermore, another technology which generated relaying on CC is “Mobile Cloud Computing” (MCC). MCC based on the concept of the “Cloud” provides any type of information and data by no matter of where and when through mobile devices. Through this relative technology the owners of the data on the internet could manage information everywhere and at any time. Also, MCC could make the mobile devices resourceful in terms such as computational power, memory, storage and energy. Considering this, MCC technology, and furthermore CC technology in general, could be settled as a base technology to operate other technologies such as BD and SNg [4] [5] [6].

A way in which the issues of data security and data privacy in SNs could be solved or could be depleted by the use of “Big Data Analysis Tools and Services”. The big data describes the data sets that are large or complex for the traditional data processing applications which are incompetent. “Big Data is often related to the use of predictive analytics or a set of advanced methods (Big Data Analytics) with the aim to extract merit from the collected data” [7] [8]. From this scope it is perceptible that the big data are now equally important both for business and internet. This happens because more data packets demand a more accurate analysis. Data analysis is a do-or-die requirement for today's businesses. The vendor community is responding by providing highly distributed architectures and new levels of memory and processing power [9] [10] [11].

The rest of the paper is divided in sections as follows. In section 2 there is a review of the related research which deals with the technology of BD and Social Networking. Section 3 discusses in detail the technology of SNg and some of its basic characteristics about its security and privacy issues. Moreover, section 4 presents and analyzes the BD technology, and some basic information about its functionality. In

Section 5, the proposed method of the paper is presented and some useful information related. Section 6, presents the proposed system-framework-network. Finally section 7 provides the conclusions of the current paper, and sets the issues of future work.

II. Related Research Review

In this section, previous studies in the field of Big Data and SNg are presented , and also the appropriate related work is examined . Below presented are the papers we have studied with their main objective.

To begin with, there are various researches for the Big Data technology. A theoretical-based review for the Big Data and the technologies that are directly connected to these amounts of data, such as Cloud Computing (CC) and Hadoop, has been presented in [12]. Also, [12] focuses on the five phases of the value chain of big data technology. As an outcome, the several representative applications of Big Data technology are examined. Furthermore, in [13] the important concepts of the big data technology are highlighted. There is also a discussion about the various features of Big Data. Furthermore, the authors of [13] define what Big Data are, and present the various parameters of its definition. Finally, in [13] there is a look at the process involved in the data processing and the security aspects of BD are reviewed. As a result, a new system for security of Big Data is proposed. Also, an establishment to the MIS Quarterly Special Issue on Business Intelligence Research is presented, which first offers “a framework that identifies the evolution, the applications and the emerging research areas of BI&A” [14]. Moreover, a definition and description of BI&A 1.0, BI&A 2.0 and BI&A 3.0 in terms of their key characteristics and capabilities are presented in [14]. Also, there exists a report of a biometric study of critical BI&A publications, researchers and research topics which rely on more than a decade of related academic and industry publications as presented in [14]. Additionally, an offer of six provocations to spark conversations about the issue of Dig Data technology is shown in [15]. Finally, a multi-stakeholder approach for developing an appropriate privacy regulation in the age of BD is presented in [16]. This argument was developed in five steps: 1) A review of the current academic debate on privacy regulation. 2) An argue that the framework for developing an appropriate privacy rule should not only focus on formal and procedural but should also consider some important substantial aspects to protect users and promote socially beneficial BD applications. 3) An examination of how the process leading to an appropriate regulation might be organized. 4) A discussion of the potential structure of a privacy organization that might conduct multistakeholder-dialogues as a precursory step. 5) A discussion of their findings and suggestions.

As follows, certain studies and analyzes have been made on the SNg field. Initially, the goal of [17] is to deploy a research model, with security and privacy concerns conceptualized as an antecedent of trust in SNg site and moderator of information sharing. Furthermore, the [17] aims to comprehend the impact of security, trust and privacy concerning the willingness of sharing information in SNg sites.

Moreover, in [18] the privacy risks associated with SNg APIs with the presentation are addressed as well as privacy-by-proxy design for a privacy-preserving API that is driven by an analysis of data-needs and data-uses of applications such as Facebook. Also, a sample for the design of a SNg privacy wizard is proposed in [19]. As an example of this specific general framework, a wizard based on an active learning paradigm called uncertainty Sampling was built. Moreover, with the aim to estimate this approach, the authors collected detailed privacy preference data from 45 real Facebook users. As a result of all these, in [19] two important things have been disclosed: 1) Real users tend to conceive their privacy preferences in terms of communities, which can easily be extracted from a SN graph using existing techniques. 2) The authors' active learning wizard, using communities as characteristics, is able to prescribe high-accuracy privacy settings using less user input than existing policy-specification tools. In addition, a description about the characteristics of SNSs, and also a proposal of a comprehensive definition presented in [20]. In addition, in [20] one aspect on the history of SN sites (SNSs) is presented, where developments and key changes have been discussed. At the end, a presentation of several of privacy and security issues, along with a design and an implementation of solutions was shown in [21]. This work lets location-based services to inquire local mobile devices for users' SN information, without disclosing user's identity or compromising users' privacy and security.

III. Security & Privacy for SNg

Social Networks can be described as web applications that permit users to create their semi-public profile [22] [23]. Most people join SNs to dispense their data and keep in contact with people that they are aware with. The main feature of SNs is a friend finder that allows SN users to search for people that they know and then build up their own online community [24].

Most SN users share a big amount of their private information in their social network space. A large number of users share their information publicly without careful consideration. Consenquently, SNs have become a large set of sensitive data. Moreover, SN users tend to have a high level of trust toward other SN users. They tend to accept friend requests easily, and trust items that friends send to them [25] [26].

Privacy and security issues on SNs are the most popular problems. The web-sites usually suffer from such problems. Meanwhile, security and privacy issues are entirely different problems. On the one hand, security issues occur when hackers gain unauthorized access to a site's protected coding or written language. On the other hand, privacy issues, those involving the unwarranted access of private information, do not necessarily have to involve security breaches. Confidential information such as typing a password can be revealed to anyone. But both types of breaches are often intertwined on SNs, especially "*since anyone who breaches a site's security network opens the door to easy access to private information belonging to any user*" [27] [28].

A. Social Networking Third-Party Output

Simple solutions are proposed for providing privacy when a SN uses third-party output. By these solutions personal data can be protected, but third party applications need direct access to the social graph information embodied in the user's friend list. More specifically, the solutions can be separated in three categories [18]: 1) Data Hiding, 2) User Identification, 3) Public Data.

IV. Big Data

BD is a more complicated world because the scale is much larger. The information is usually shared over a number of servers, and the work of compiling the data must be correlated among them. In the past, the work was largely delegated to the database software, which would use its magical JOIN [29] mechanism to compile tables, then add up the columns before handing off the rectangle of data to the reporting software that would paginate it. Database programmers can inform the users about the procedure about complicated JOIN commands that would lock up their database for hours as it tried to produce a report for the boss who wanted his columns just so [29] [30] .

BD sets advanced analytic techniques in which they operate on, that called BD Analytics. Therefore, BD analytics is about two things, BD and analytics, plus how the two have teamed up to produce one of the most profound trends in business intelligence (BI) today. Analytics helps us discover what has changed and how we should react [31] [32].

A. BD Features

Most definitions of BD focus on the size of data in storage. Size matters, but there are other important attributes of BD, namely data variety and data velocity. The three Vs of BD, which are volume, variety, and velocity, constitute a broad definition, and they bust the myth that BD is only about data volume. More specifically, each one of these three Vs has its own ramifications for analytics [31].

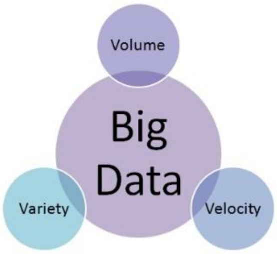


Figure 2. The Three Vs of Big Data.

1) Big Data Volume

2) Big Data Velocity

3) Big Data Variety

B. BD Analysis Tools and Services

BD is the emerging discipline of capturing, storing, processing, analysing and visualising these huge quantities of information. The data sets may start at a few terabytes and run to many petabytes, far more than traditional data analysis packages can handle [33] [34].

Some BD tools that analyzed bellow are: 1) Jaspersoft BI Suite, 2) Pentaho Business Analytics, 3) Karmasphere Studio and Analyst, 4) Talend Open Studio, 5) Skytree Server, 6) Tableau Desktop and Server, 7) Splunk.

C. Big Data's impact in SNg

As the BD technology grows and spreads on the internet, many web technologies and applications that rely on it are affected. One of the many applications which are affected by the growth of the BD technology is the SNg.

TABLE I.
Big Data's characteristics affect on Social Networking's third party output**.**

Big Data Characteristics	**Big Data Volume**	**Big Data Velocity**	**Big Data Variety**
Data Hiding	**X**	**X**	**X**
User Identification	**X**	**X**	
Public Data	**X**		**X**

Table 1 lists the characteristics that the BD technology has, regarding the convenience that this technology offers, and on the other hand lists three categories of the SNg third party output. The aim of the Table 1 is to show how the characteristics of the BD technology are related to the the three categories of Big SNg third party output, and additionally how they affect these three categories. The conclusion that can be drawn from Table 1 is that the BD Volume affects more in the SNg third party output. We reach to this conclusion relying to our study on Big Data technology, and in addition the findings and the conclusions of the related works, which we have studied.

TABLE II.

Social Networking's third party output categories affect on Big Data's Analysis Tools & Services**.**

Big Data Analysis Tools & Services	**Dat a Hiding**	**User Identification**	**Public Data**
Jaspersoft BI Suite		**X**	**X**
Pentaho Business Analytics	**X**		**X**
Karmasphere Studio and Analyst			**X**
Talend Open Studio		**X**	**X**
Skytree Server	**X**	**X**	**X**
Tableau Desktop and Server	**X**	**X**	
Splunk		**X**	**X**

Table 2 lists three categories of the SNg third party output and on the other hand lists the Big Data's Analysis Tools & Services that we have studied in this paper. The aim of Table 2 is to show how the three categories of the SNg third party output related and affect the Big Data's Analysis Tools & Services. As shown, Table 2 demonstrates that the Public Data category are related more with the Big Data's Analysis Tools & Services which we have studied here. Also, another conclusion that can be drawn from Table 2 is that the Skytree Server was affected more by the three categories of the SNg third party output.

D. BD Security Issues

New challenges and standards developed and created in data security issues through the development and the use of BD technology. This creates a growing need for further research on security technologies in order to make handling the huge amount of data feasible and to ensure effectiveness. Technologies for securing data are slow when applied to huge amounts of data [35].

TABLE III.

Encryption Rates of popular Algorithms.

Algorithm	Key length	MB processed	Block size	Rounds	Time Taken	MB per Second
Blowfish	32-448 bits	256	64 bits	16	3,976	64,386
DES	56 bits	128	64 bits	16	5,998	21,340
3DES	56, 112 or 168 bits	128	64 bits	48	6,159	20,783
AES	128, 192 or 256 bits	256	128 bits	10, 12 or 14	4,196	61,010
RSA	1024-4096 bits	300	512 bits	1	1175,7826	10,900

Regarding Table 3, the conclusion that even the most efficient algorithms give an encryption rate of 64.3MB/s is reached. So, in the sector of BD technology, in which the need of large amounts of data to be transferred, we can confirm a significant bottle neck for encryption such large amounts data. This is detrimental to the nature of BD which have real time processing and results.

V. Evaluation Experiments

As the BD technology develops and engages with other technologies, established new requirements result relating to operation and needs. Thus there exists a causality of BD technology with an equally growing technology over the last years, which is the the Social Networking.

Having studied some encryption algorithms regarding security issues of BD technology we find that with regard to security issues involving BD technology in SNg technology, there are some issues which can be combined in a Cloud Environment. Selecting two of the encryption algorithms that were previously studied, we attempt to modify them so they can be use data from the algorithms we use in the SNg technology with the aim to realize some specific measurements of the data can be obtained, and why not do it in a safer way. The two algorithms are selected based on their potential to receive more data per second. The algorithms are the Blowfish (64,386MB/s) and the AES (61,010MB/s).

Regard the Blowfish algorithm we can take the NIter, which is the maximum number of iterations with the aim to use it in the encryption algorithm, and to improve the security of the four different bio-inspired algorithms.

void encrypt (NIter & L, NIter & R) {…}

void decrypt (NIter & L, NIter & R) {…}

Regarding the AES algorithm we can take the same value, the NIter, which is the maximum number of iterations in order to use it in the encryption algorithm, and to improve the security of the four different bio-inspired algorithms.

*int mbedtls_aes_crypt_ecb(NIter *ctx, int mode, const unsigned char input[16], unsigned char output[16]) {…}*

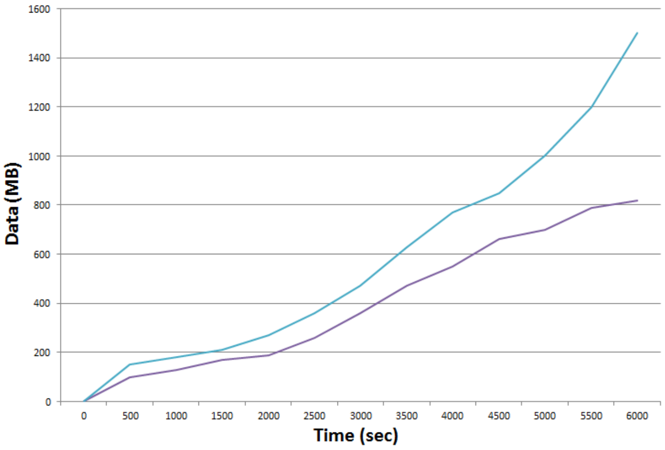


Figure 3. Security level of encryption algorithms of measurement used for the study of four bio-inspired algorithms.

Figure 3 shows, the measurements with respect to time. As can be deduced by this figure the more often the combined use of the algorithms, the higher the level of security of the data we achieve every time. The upper line represents the Blowfish algorithm and the other (down line) represents the AES algorithm. More specific, Figure 3 could demonstrate a comparison of the implementation and the use of the two aforementioned encryption algorithms. The graph represents that through time the encryption procedure become more accurate and more efficient.

Regarding the Encryption Rate of the Transmitted Data the following equation is considered: the related work we derive the following equation:

$$E_nR_a = \frac{\mathrm{R_e}\,D_{ata} - \left(T_rD_{ata} * NI_{ter}\right)}{T_rT_{ime}} \qquad (1)$$

where,

Acronym	Description
E_nR_a	Encryption Rate of Data
R_eD_{ata}	Received Data
T_rD_{ata}	Transmitted Data
NI_{ter}	Maximum number of iterations
T_rT_{ime}	Transmission Time of Data sent

By applying, so equation (1), the following chart occurs.

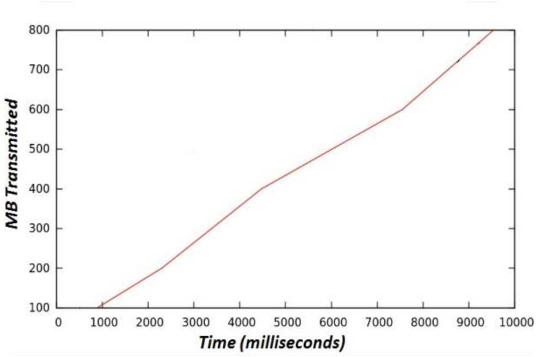


Figure 4. Encryption Rate (*Data in Time*).

As observed from Figure 4, the encryption rate has an upward trend over time, in respect of data transmitted on the network. So, we can conclude that we need a good implementation of the encryption process mainly before sending the data, and then later in the transmission process.

Counting on the data packet switching procedures and use of the data on the internet, you need to make a further study of additional technologies, such as Internet of Things (IoT), in order to see if combined we can achieve better results on data usage and security issues [36] [37].

VI. Proposed System

Considering the study conducted for the related review, we can conclude that creating a system-framework-network in a “safe” Cloud environment through which users of the various Social Networks will be able to exchange data and information, and primarily large-scale data (Big Data), could greatly improve the communication of SN users.

In addition, having studied the available ways of security and authentication offered by social network providers, we reached that the system that we propose should work with authentication (sign in) through the account that will every user have in a SN (e.g. Facebook, LinkedIn YouTube, Instagram). In this way, each user will be able to connect to a more secure "private" network through which the user can exchange data in a “*Safe Cloud Server*” with other SN users, such as photos (mainly high quality) and videos (mostly high quality).

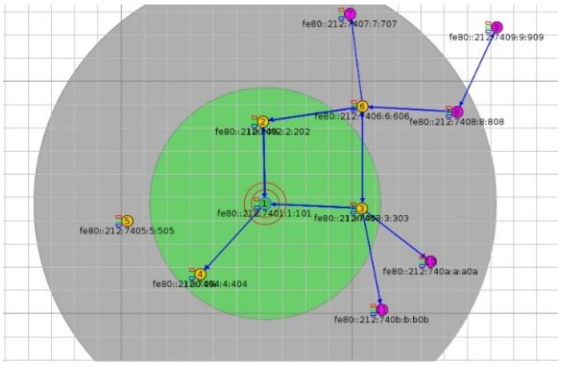
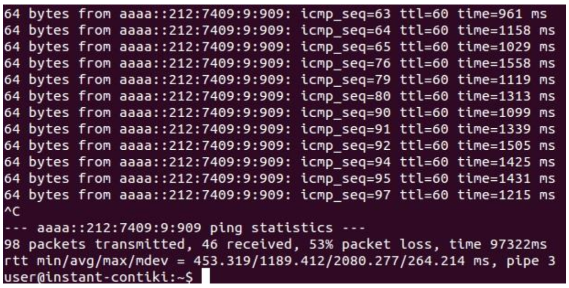


Figure 5. Propoesd System’s Topology.

Having concluded that many similar technologies in the telecommunications sector can be combined with each other from the earlier study we have done, the network we created will be based, in terms of design, architecture, topology, on IoT technology. This type of network that we propose could be count on previous work of C. Stergiou et al [38], where a new type of network topology have been proposed (figure 5) in order to transmit high quality videos. Also, users of this network will be able to exchange and other types of data such as personal files, which can be quite large. Users of the network will also be able to temporarily store data, as well as back up data, during the transmission process in a network space, based on CC technology. For multimedia data (Big Data) transfer within the network a protocol that has been proposed in a previous work of G. Kokkonis et al [39] will be used, the NAMRTP.

The proposed network system will use existing models of cryptographic algorithms to secure the authentication and data exchange process. Of course, there are some improvements - changes to some pieces of their source code, as we have seen in the previous section. The aim concerning the network will be to offer an alternative and more secure data exchange solution among users of SNs.



```
64 bytes from aaaa::212:7409:9:909: icmp_seq=63 ttl=60 time=961 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=64 ttl=60 time=1158 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=65 ttl=60 time=1029 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=76 ttl=60 time=1558 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=79 ttl=60 time=1119 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=80 ttl=60 time=1313 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=90 ttl=60 time=1099 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=91 ttl=60 time=1339 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=92 ttl=60 time=1505 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=94 ttl=60 time=1425 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=95 ttl=60 time=1431 ms
64 bytes from aaaa::212:7409:9:909: icmp_seq=97 ttl=60 time=1215 ms
^C
--- aaaa::212:7409:9:909 ping statistics ---
98 packets transmitted, 46 received, 53% packet loss, time 97322ms
rtt min/avg/max/mdev = 453.319/1189.412/2080.277/264.214 ms, pipe 3
user@instant-contiki:~$
```

Figure 6. Packets send throug Network (sample node 9).

Figure 6 shows the transmission procedure of packets sent through the network. As easily observed, each file that ends through the proposed network is divided to smaller packets of data in order to be sent. Regarding the large amount of data sent we have a small number of *Packet Loss*. More specifically, the node 9 which is shown in figure 6 is the most distant node of the simulation network.

$$P_a L_o = \frac{\left(P_a T_r - P_a R_e\right) - D_u P_a}{T_r T_{ime}} \qquad (2)$$

$$\frac{P_a R_e}{T_r T_{ime}} = \frac{P_a T_r - P_a L_o - D_u P_a}{T_r T_{ime}} \qquad (3)$$

where,

Acronym	Description
P_aL_o	Packets Loss
P_aT_r	Packets Transmitted
P_aR_e	Packets Received
D_uP_a	Duplicated Packets
T_rT_{ime}	Transmission Time of Packets sent

The (2) shows the *Packet Loss* of the transmission procedure through the proposed network. The rate of the *Packet Loss* differs through time and depends by the various amount of data send each time. While, on the other hand, (3) shows how the *Packages Received* during the Transmission process (Time) depend, from the *Total Packets Transmitted*, removing the *Packet Loss* and the *Duplicated Packets*, and dividing them by the *Transmission Time*.

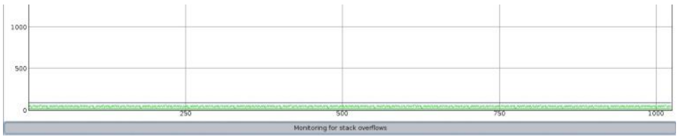
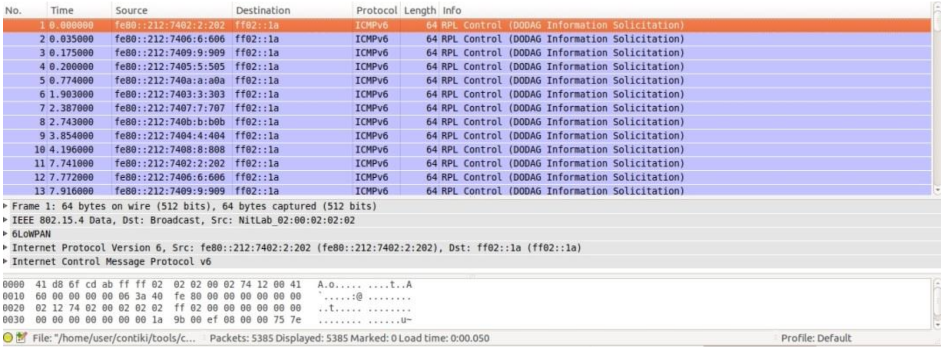


Figure 7. Stuck overflow not detected.

Figure 7 demonstrates that there is no stuck overflow during the transmission procedure, so we can deduce that the whole process is smoothly carried out in the network.



No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
2	0.035000	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
3	0.175000	fe80::212:7409:9:909	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
4	0.200000	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
5	0.774000	fe80::212:740a:a:a0a	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
6	1.903000	fe80::212:7403:3:303	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
7	2.387000	fe80::212:7407:7:707	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
8	2.743000	fe80::212:740b:b:b0b	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
9	3.854000	fe80::212:7404:4:404	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
10	4.196000	fe80::212:7408:8:808	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
11	7.741000	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
12	7.772000	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
13	7.916000	fe80::212:7409:9:909	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)

```
▸ Frame 1: 64 bytes on wire (512 bits), 64 bytes captured (512 bits)
▸ IEEE 802.15.4 Data, Dst: Broadcast, Src: NitLab_02:00:02:02:02
▸ 6LoWPAN
▸ Internet Protocol Version 6, Src: fe80::212:7402:2:202 (fe80::212:7402:2:202), Dst: ff02::1a (ff02::1a)
▸ Internet Control Message Protocol v6

0000  41 d8 6f cd ab ff ff 02  02 02 00 02 74 12 00 41   A.o..... ....t..A
0010  60 00 00 00 00 06 3a 40  fe 80 00 00 00 00 00 00   `.....:@ ........
0020  02 12 74 02 00 02 02 02  ff 02 00 00 00 00 00 00   ..t..... ........
0030  00 00 00 00 00 00 00 1a  9b 00 ef 08 00 00 75 7e   ........ ......u~
```

File: "/home/user/contiki/tools/c... Packets: 5385 Displayed: 5385 Marked: 0 Load time: 0:00.050 Profile: Default

Figure 8. Transmission proccess (a).

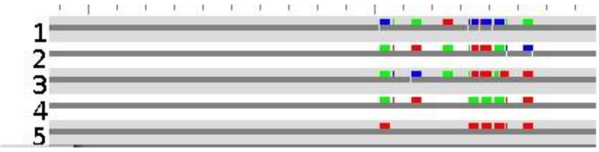


Figure 9. Transmission proccess (b).

Figures 8 and 9 show the Transmission Packets procedure through the network. As observed, various types of packets were sent in the network by a number of connected users.

VII. Conclusions

SNg technology offers many possibilities, but also places several limitations as well. Social Networking could be described as web applications that allow users with the aim to create their semi-public profile. In the current work, we survey SNg, BD and Cloud Computing (CC) technology and their basic characteristics, with a focus on the security issues of those technologies. Additionally, we presented the basic characteristics of BD an SNg technologies, and also the major privacy and security issues that both technologies face. Subsequently and in terms of BD technology, we survey the algorithms with big impact to its security, and we present the basic characteristics of them.

Finally, we discuss the opportunity to create a database through which each user can see the statistics of his interaction with the SNg. The main goal of this paper is to try to combine the functionality of the BD and SNg technologies in a CC environment, in order to examine the common features, and also to discover the benefits related in security issues of their integration. Also, by examining their integration and functionality we could establish a new system-framework-network in Cloud Environment that combines these technologies, and some other technologies (e.g. IoT) related. This could be take place by presenting a new system-framework-network through which users of the various Social Networks will be able to exchange data and information, and primarily large-scale data (Big Data) and greatly improve the communication of SN users, and thus become more safe and accurate in a Cloud environment. Meanwhile, this system could be used for the purpose of improving security of SNg users through the use of algorithms that can provide more privacy in the data related to BD technology in a Cloud Server. This method is presented here and also some measurements results of its use.

This can be a field of future research on the integration of those technologies, and also have a huge improvement of their security and privacy issues. In addition to this, we can conclude that it would be a useful opportunity to create a database through which each user can see the statistics of his interaction with the SNg. Furthermore, based on the rapid development of network technologies the plethora of new technologies in this field, it would be good a further study to consider related technologies such as IoT, as a new case study.

VIII. References

[1] S. Wasserman, K. Faust, “Social Network Analysis: Methods and Applications”, Urbana-Champaign: Cambridge University Press, pp. 1-27, March 1995.

[2] M. Newman, A.-L. Barabasi, D. J. Watts, “The Structure and Dynamics of Networks”, ACM, Princeton University Press Princeton, NJ, USA, 2006.

[3] C. Fabiana, M. Garetto, E. Leonardi, “De-anonymizing scale-free social networks by percolation graph matching”, in Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM), Kowloon, Hong Kong, 26 April-1 May 2015.

[4] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, December 2016.

[5] S. Sakr, A. Liu, D. M. Batista, & M. Alomari, “A survey of large scale data management approaches in cloud environments”, IEEE Commun. Surveys & Tutorials, vol. 13, no. 3, pp. 311–336, 2011.

[6] C. Stergiou, K. E. Psannis, "Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey", Wiley Online Library, International Journal of Network Management, vol. 27, issue 3, pp. 1-12, May 2016.

[7] M. Hilbert, P. Lopez, "The World’s Technological Capacity to Store, Communicate, and Compute Information", Science, vol. 332, issue: 6025, pp. 60-65, April 2011.

[8] Z. Fu, K. Ren, J. Shu, X. Sun, F. Huang, "Enabling Personalized Search over Encrypted Outsourced Data with Efficiency Improvement", IEEE Transactions on Parallel and Distributed Systems, vol. 27, issue: 9, September 2016.

[9] W. Culhane, K. Kogan, C. Jayalath, P. Eugster, “Optimal communication structures for big data aggregation”, in Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM) , Kowloon, Hong Kong, 26 April-1 May 2015.

[10] A. Detsounis, G. S. Paraschos, I. Koutsopoulos, “Streaming big data meets backpressure in distributed network computation”, Computer Communications, in Proceedings of 35th Annual IEEE International Conference on IEEE INFOCOM 2016, San Francisco, CA, USA, 10-14 April 2016.

[11] Z. Su, Q. Xu, Q. Qi, “Big Data in Mobile Social Networks: A Qof-Oriented Framework”, IEEE Network, February 2016.

[12] I. Chebbi, W. Boulila, I, R, Farah, "Big Data: Concepts, Challenges and Applications", Springer International Publishing, Computational Collective Intelligence, vol. 9390, pp. 638–647, October 2015.

[13] R. Toshniwal, K. G. Dastidar, A. Nath, "Big Data Security Issues and Challenges", International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 2, issue: 2, pp. 15-20, February 2015.

[14] H. Chen, R. H. L. Chiang, V. C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact", ACM, MIS Quarterly, vol. 36, issue: 4, pp. 1165-1188, December 2012.

[15] D. Boyd, K. Crawford, "Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon", Information, Communication & Society Journal, vol. 15, issue: 5, pp. 662-679, May 2012.

[16] M. G. Will, “Privacy and Big Data: The Need for a Multi-Stakeholder Approach for Developing an Appropriate Privacy Regulation in the Age of Big Data”, Halle (Saale) Universität Halle-Wittenberg, SSRN Electronic Journal, vol. 3, July 2015.

[17] A. Dhami, N. Agarwai, T. K. Chakraborty, B. P. Singh, J. Minj, "Impact of trust, security and privacy concerns in social networking: An exploratory study to understand the pattern of information revelation in Facebook", in Proceedings of 3rd IEEE 2013 International Advance Computing Conference (IACC), Ghaziabad, India, 22-23 February 2013.

[18] A. Felt, D. Evans, "Privacy Protection for Social Networking APIs", W2SP, Workshop on Web 2.0 Security and Privacy, Oakland, CA, May 2008.

[19] L. Fang, K. LeFevre, "Privacy Wizards for Social Networking Sites", in Proceedings of the 19th international conference on World wide web WWW '10, pp. 351-360, NC, USA, 26-30 April 2010.

[20] D. M. Boyd, N. B. Ellison, “Social Network Sites: Definition, History, and Scholarship”, Wiley Online Library, Journal of Computer-Mediated Communication, vol. 13, issue: 1, pp. 210–230, October 2007.

[21] A. Beach, M. Gartrell, R. Han, "Solutions to Security and Privacy Issues in Mobile Social Networking", in Proceedings of International Conference on Computational Science and Engineering, 2009. CSE '09, Vancouver, BC, Canada, 29-31 August 2009.

[22] T. Ma, J. Zhou, M. Tang, S. Lee, "Social network and tag sources based augmenting collaborative recommender system", IEICE Transactions on Information and Systems, vol. E98-D, no.4, pp. 902-910, April 2015.

[23] L. Bilge, T. Strufe, D. Balzarotti, E. Kirda, "All your contacts are belong to us: automated identity theft attacks on social networks", in Proceedings of the 18th international conference on World Wide Web WWW '09, pp. 551-560, Madrid, Spain. 20-2 April 2009.

[24] J. L. Z. Cai, M. Yan, Y. Li., “Using crowdsourced data in location-based social networks to explore influence maximization”, in Proceedings of the 35th Annual IEEE International Conference on Computer Communications IEEE INFOCOM 2016, San Francisco, CA, USA , 10-14 April 2016.

[25] P. Chaudhary, B. B. Gupta, S. Gupta, "Auditing Defense against XSS Worms in Online Social Network-Based Web Applications," Handbook of Research on Modern Cryptographic Solutions for Computer and Cyber Security, IGI-Global's Advances in Information Security, Privacy, and Ethics (AISPE) series, USA, 2016.

[26] J. Cheng, Y. Zhang, Q. Ye, H. Du, “High-precision shortest distance estimation for large-scale social networks”, in Proceedings of the 35th Annual IEEE International Conference on Computer Communications, IEEE INFOCOM 2016, San Francisco, CA, USA, 10-14 April 2016.

[27] D. Gunatilaka, "A Survey of Privacy and Security Issues in Social Networks," CSE571S: Network Security, pp. 1-12, November 2011.

[28] L.Yan, H. Shen, K. Chen, “TSearch: Target-oriented low-delay node searching in DTNs with social network properties”, in Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM) , Kowloon, Hong Kong, 26 April-1 May 2015.

[29] P. Wayner, "7 top tools for taming big data," InfoWorld, 18/4/2012. [Online]. Available: http://www.infoworld.com/article/2616959/big-data/7-top-tools-for-taming-big-data.html. [Accessed 21/5/2016].

[30] A. P. Plageras, C. Stergiou, G. Kokkonis, K. E. Psannis, Y. Ishibashi, B.-G. Kim, B. B. Gupta, “Efficient Large-scale Medical Data (eHealth Big Data) Analytics in Internet of Things”, in Proceedings of 2017 IEEE 19th Conference on Business Informatics (CBI), International Workshop on Internet of Things and Smart, Thessaloniki, Greece, 24-26 July, 2017.

[31] Cloud News Daily, "Guide to Big Data Analytics: Platforms, Software, Companies Tools, Solutions and Hadoop," Cloud News Daily, 12/12/2015. [Online]. Available: http://cloudnewsdaily.com/big-data-analytics/. [Accessed 21/5/2016].

[32] K. Zheng, Z. Yang, K. Zhang, P. Chatzimisios, K. Yang, W. Xiang, "Big Data-Driven Optimization for Mobile Networks toward 5G", IEEE Network, February 2016.

[33] S. Kaisler, F. Armour, J. A. Espinosa, W. Money, "Big Data: Issues and Challenges Moving Forward", in Proceedings of 2013 46th Hawaii International Conference on System Sciences (HICSS), pp. 995-1004, Wailea, Maui, HI, USA, 7-10 January 2013.

[34] K. Raichura, N. Padhariya, "BigCache: a cache-based Big Data management in mobile networks", International Journal in Mobile Communications, vol. 15, no. 1, pp. 49-68, 2017.

[35] C. Stergiou, K. E. Psannis, "Algorithms for Big Data in Advanced Communication Systems and Cloud Computing", in Proceedings of 2017 IEEE 19th Conference on Business Informatics (CBI 2017), Thessaloniki, Greece, 24-26 July 2017.

[36] C. Stergiou, K. E. Psannis, "Efficient and Secure Big Data delivery in Cloud Computing", Springer, Multimedia Tools and Applications, vol. 76, issue 21, pp. 22803–22822, November 2017.

[37] K. Yang, X. Jia, K. Ren, R. Xie, L. Huang, "Enabling efficient access control with dynamic policy updating for big data in the cloud", in Proceedings of 2014 IEEE INFOCOM, Toronto, ON, Canada, 27 April-2 May 2014.

[38] C. Stergiou, K. E. Psannis, A. P. Plageras, Y. Ishibashi, B. B. Gupta, B.-G. Kim, "Architecture for security monitoring in IoT environments", in Proceedings of 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, Scotland (UK), 19-21 June 2017.

[39] G. Kokkonis, K. E. Psannis, M. Roumeliotis, D. Schonfeld, "Real-time wireless multisensory smart surveillance with 3D-HEVC streams for internet-of-things (IoT)", Springer, Journal of Supercomputing, vol. 73, issue: 3, pp. 1044-1062, March 2017.

3. Work Title: Proposed High Level Architecture of a Smart Interconnected Interactive Classroom

C. Stergiou, A. P. Plageras, K. E. Psannis, T. Xifilidis, G. Kokkonis, S. Kontogiannis, K. Tsarava, A. Sapountzi, "Proposed High Level Architecture of a Smart Interconnected Interactive Classroom", in Proceedings of IEEE conference SEEDA-CECNSM 2018, 22-24 September 2018, Kastoria, Greece.

Abstract

This paper presents the high level architecture of a smart, modern, interactive laboratory-class called Smart Interconnected Interactive Classroom (SIIC). It describes the interoperability of telecommunication technologies, sensors and actuators over a virtual environment that enhances the learning process and experience. In the context of this work novel augmented and virtual services are outlined that can assist e-Learning systems through virtual reality and real-time interactions.

Keywords - Smart Classroom; Haptic technology; Cloud services; Internet of Things; Big Data; middleware protocols; virtual reality

I. Introduction

Educational Learning Management systems (LMS), are of high impact in terms of technology appliance and testing methodologies. Although the field of smart-Education has been scientifically established, it is currently in an embryonic 2D data representational state. Contemporary LMS systems include components such as Forum, Wiki, knowledge surveys, tasks, document management, games, reporting that assist teaching process. These aggregations of services are part of the asynchronous LMS functionality, while the use of VoD services, and real-time, mobile audio-video course conducting services are part of the synchronous LMS functionality [13-15]. In this paper the authors propose the incorporation of 3D virtual services in an LMS platform that will include both synchronous and asynchronous services into the virtual class. Furthermore, the proposed virtual class will utilize sensors and haptic equipment [11] together with state-of-the-art hardware implemetation and sensors in order to carry out augmented human sensing information and touch into the virtual class. Hence, based on the above, the application of technologies mentioned for the implementation of the proposed Smart Interconnected Interactive Classroom (SIIC) constitutes an ambitious step and a perspective that will inaugurate progress in the first and second grade education [8-10]. The proposal of this paper must be fully aligned to software solutions accompanied by advanced hardware solutions that efficiently support the services provided. The proposed SIIC architecture is presented at section II. At section III, the authors outline proposed novel services and protocols that will embrace virtuality, while at section IV, the authors present considerations and implementation plan.

Internet of Things (IoT), Big Data (BD) and cloud technologies are already well-established and have progressed rapidly, counting many years of life and scientific interest [1, 3]. In the Cloud Services (CS) field, data compression and delay tolerant data representation is relatively more recently spread out because of the IoT technological outburst.

Regarding the aforementioned technologies, starting with Cloud Computing (CC), it is consisted as a technology of internet services providing remote use of hardware and software. Thus, the users of CC could have access to information and data from any place at any time. CC in general, could be settled as a base technology to operate other technologies such as Internet of Things and Big Data. Moreover, we could realize that the basic idea of the IoT is the pervasive presence of a variety of things or objects used by people such as radio-frequency identification tags, sensors, actuators, and mobile phones. Finally, as regards the Big Data, we could define that it is a new popular term, used to describe the surprisingly rapid increase in volume of data in structured and unstructured form. BD usually uses Cloud Computing (CC) as a base technology in order to operate [1-3].

Wireless communications include technologies and equipment for data collection from wireless sensors, such as temperature, smoke, humidity, capacitive touch, and task instructed communication protocols such as streaming, real-time, interactive, responsive and best effort. At this point, it should be noted that the sensor network efficiency must be carefully addressed as to ensure reliable data gathering. Hence, the network topology must comply to the location of the indoor-smart classroom user terminals and the main factors of signal degradation and attenuation. This aggregation of devices and specialized protocols connected to the Internet cloud and focused only to a specific user is part of the immersing revolution of cloud-oriented, multi-disciplinary user targeted services, named after the nickname smart [2]. It is a fact that concerning that state-of-the-art technology Greece is not highly ranked. Thus, as the transition to the state-of-the-art 5G networks is realized and specifically considering indoor application, the convergence of the aforementioned technologies combined with smart technology is at the center of scientific interest [5-7]. Moreover, contemporary BLE, LoRa and XBee low power networking technologies for sensory data acquisition are still not fully exploited. In addition, obsolete technologies of 2G/3G and SCADA-RTU/smart equipment communication protocols (DNP3, Modbus, Ethernet IP based IEC 61850, IEC60870, RS232/Parport EPICS). Even SCADA systems nowadays, with the development of IoT and cloud technology are moving towards that adoption.

Virtual reality is a known scientific area of high interest. Combined with the Haptic sense and haptic protocols for haptic data transfer via Internet, the proposed system will provide capabilities of conducting experiments in courses such as Physics and using on-line applications in an interactive distance learning [11, 12]. The students will have the capability of developing active learning and improving their knowledge level, while at the same time motivation will be provided from the teacher as well as the interconnected interactive technology itself, thus contributing to the overall improvement of their educational and technological training level. Moreover, the students, being part of this virtual smart interactive classroom will gain better understanding of the concepts presented in the courses and develop useful skills and abilities of solving complex problems. The appliance of the proposed architecture in every school in first and second grade education will have a significant impact as it will improve education quality with long-term benefits in the educational level and the scientific training of young scientists.

II. Literature Review

For the purpose of this work we count on previous literature works which has been published in the related field. The following paragraphs present the papers which contributed significantly in our study.

In [1] the authors survey BD and CC technology and their basic characteristics, with a focus on the security and privacy issues of both technologies. Particularly, the authors try to combine the functionality of the two technologies with the aim to examine the frequent features, and also to discover the benefits related in security issues of their integration. Additionally, this work presents a new method of an algorithm that can be used for the purpose of improving Cloud Computing’s security through the use of algorithms that can provide more privacy in the data related to BD technology.

The [3] presents a survey of IoT and Cloud Computing with a focus on the security issues of both technologies. More specifically, the authors try to combine these two technologies with the aim to examine the common features, and in order to discover the benefits of their integration. Through this work, it is shown how the Cloud Computing technology improves the function of the IoT.

The [4] presents related work on High-Efficiency Video Coding. It points out the challenges and the synchronization techniques that have been proposed for synchronizing video and haptic data. Resulting, the [4] proposes a new efficient algorithm for transferring a real-time HEVC stream with haptic data through the Internet.

The authors of [9] try to attempt the evaluation of an educational scenario, where the implementation of which is based on Cloud Computing tools that serve collaborative learning. The aim of the collaborative activities of the script is to understand and consolidate the usefulness of the criteria that make an educational video appropriate or not, for its introduction into the educational process.

The [14] describes the use of data mining techniques, such as clustering, classification, and association, in order to analyze the log file of an e-Learning platform and deduce useful conclusions. Also, a case study based on a previous approach was applied to e-Learning data from a Greek University.

III. SIIC High Level System Architecture

The authors propose a novel system architecture over a virtual reality world (realm) and define the services that will support user-real interaction. This virtual architecture will support a virtual lab environment and it is called as Smart Interconnected Interactive Class (SIIC). The main objective of the proposition is to enhance learning process out of the class boundaries, supporting user sense transfer, and distant interaction. Smart Interconnected Interactive Classroom is consisted of the following structural parts: a) Interactive interface workstations, equipped with Haptic equipment and sensors b) The architecture of each interactive interface workstation that comprises the interactive class user stations is illustrated at Figure 1.

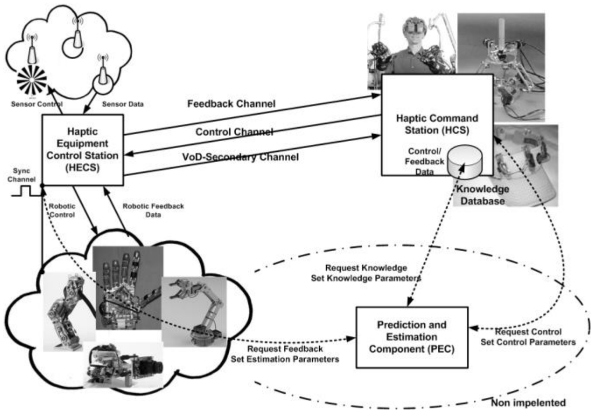

Sensor Control
Sensor Data
Feedback Channel
Control Channel
VoD-Secondary Channel
Haptic Equipment Control Station (HECS)
Haptic Command Station (HCS)
Control/ Feedback Data
Knowledge Database
Sync Channel
Robotic Control
Robotic Feedback Data
Request Knowledge Set Knowledge Parameters
Prediction and Estimation Component (PEC)
Request Control Set Control Parameters
Request Feedback Set Estimation Parameters
Non impelented

Figure 1: Student interface system-workstation architecture of a workbench providing enhanced and interactive virtual reality capabilities.

The practical implementation of such an interconnected environment can be realized with the direct-real-time networking capabilities enforcement among the instructor and the workstations equipped sensor-actuator devices (human machine interfaces) of his students. This requires the implementation of appropriate interactive and real-time protocols either at application level (for each of the aforementioned services separately) or at the interoperability of sensor systems and data transfer in the Learning Management System (LMS) [4, 17].

The proposed SIIC architecture will be comprised of the following: 1) a cloud computing server LMS (Learning Management System) which acquires periodically and real-time data streams of users, actuators and sensors, stored from a set of tactile devices and equipment inter-connected via a local wireless network. [15], 2) For the haptic devices data handling, appropriate application/data transfer protocols will be designed and implemented according to [4, 7, 12, 17] 3). Computer devices, haptic devices [11], virtual and augmented reality headsets, and student sensors will be installed in the proposed classroom in to each one workstation entity, where the testing of protocols and human-machine interoperability with the virtual world will be performed. A very promising scientific field is that of providing feedback which could significantly enhance quality of services based on the fact that future data requirements may be previously estimated based on frequently appearing patterns. The benefit is two-fold: the future network data processing could be considerably improved based on the "memory" of the whole system and energy efficiency particularly concerning sensor limited battery-life could be achieved if prediction is performed off-line.

The instructor will also have online access to the modified LMS system and equipped with virtual course capabilities and interfaces with the virtual class. The modified product will be named as LMS system of Virtual Context (LMSVC), where user sensory data will be stored in real time or interactively during the course. In front of the LMSVC system a balancing controller shall be used [16] that will monitor user interaction, besides the data usage of the devices, sensors and actuators data that will be available for data mining and knowledge extraction. The LMSVC will offer the virtual creation and the previously mentioned interactive real-time virtual services, as well as the capabilities of conventional 2D-LMS systems, such as virtual Documents, virtual announcements, virtual self-evaluation virtual exercises, virtual works and questionnaires, forum wiki, etc.

These services of conventional LMS systems will be provided offline and distant (without the requirement of course conduction in the interactive class), in the virtual learning world where the student with the virtual reality headset can be connected to the LMS Virtual Balancer [16], download documents through his/her tablet, computer and assess LMS system services in a 3D visual sense [20]. The LMS application virtual services and data transfer protocols enclosed, as well as the pilot headset application that will interface the student to the LMSVC system will be implemented within the framework of the project.

The SIIC system will enable pupils to approach the scenes of the modern virtual educational process and to get in touch with modern supervisory tools. The social impact of SIIC will be particularly important, as First and Secondary Education will have access either to a physical presence or to the Internet in this experimental class. Students will understand through empirical learning, with the help of augmented reality. All students' interaction with the system will be recorded and new teaching methods will emerge through the study of such behavior.

IV. Services of the Smart Interconnected Interactive Classroom

The authors proposed virtual services to be tested and implemented within SIIC apart from existing synchronous and asynchronous services are the following:

1. **Virtual classroom service.** This service will enable the student to immerse in an interactive three dimensional environment of the class from artificial virtual imaging projected by the LMS system to the end user using 3D virtual reality

equipment (VR-Box, VR-Glasses). The interaction between the user members (avatars) of this virtual class will be carried out from the real world to the virtual world through the use of tactile-haptic devices and appropriate 3D modeling and presentation layers. The network protocols that will support this service and the virtual illustration of existing LMS modules are the standard best-effort HTTP-TCP protocols for connection oriented services, UDP protocols for text and messaging services and RTP-RTCP protocols as well as experimental ones for HEVC streaming services [7, 4].

2. **Cognitive Service and Augmented sensory services**. Augmented reality service will utilize sensors for monitoring user body activity and bio-readings [17]. Such readings will be transferred to the virtual world and will be illustrated to the virtual class and virtual user navigation among the virtual class facilities-LMS components (augmented-sense service). Energy conservation is a crucial optimization parameter for the wireless sensor as stated previously. Based on the authors' claim that this smart interactive classroom is built with state-of-the-art components and operations the sensors could benefit from energy harvesting methods which involve collecting energy from neighboring networks. Although this could prove costly it will prolong sensor life-time.

Augmented reality service will also provide in real time via EMG sensors, temperature sensors, sweat sensors, augmented information of the student's mental and psychological state. By using sensory data and implementing artificial intelligent and data mining algorithms with appropriate pattern profiles, can offer measurements of user mood indication and user level of course understanding. Such capabilities are part of the proposed cognitive service that will enable adequate supportive information like whether the student is interested, or understands the delivered content of the virtual lesson or if the user is not in a disordered attention state along with the user's current perception and course interest [8]. Also this cognitive service real-time reporting can be recorded and used as feedback by other student-course evaluation services [14]. Augmented reality and cognitive services will use protocols that offer asynchronous communications in the context of request and response such as the CoAP protocol. In cases of sensor measurements of periodic synchronous transmissions the MQTT protocol will be used. In addition, experimental protocols for periodic and asynchronous IoT devices, proposed by the authors for medical services will replace CoAP and MQTT as more efficient [1, 2, 3, 17].

3. **Positioning service.** This service will control the student's current position within the virtual class in contrast to its real-class indoor position. For the process of indoor positioning, a set of Wi-Fi / BLE transceivers (iBeacons) – will be used with the implementation of location-positioning algorithms in the LMS system that will provide the exact position-placement of the student inside the classroom projected to the virtual world. Positioning service will use two different types of protocols for data delivery of movements to the virtual world. Using a periodic synchronous protocol with no ACK feedback such as UDP for carrying little motion information and a high resolution stream protocol such as RTP accurate positioning due to collaborative activities is required [9, 17].

4. **Touch interaction service-Haptic service.** This service will provide the ability to visualize the sense of touch in the virtual world and experience the virtual touch of others in the real world. This service will be experimentally tested with the use of specially designed gloves equipped with appropriate analogue pressure sensors

and infrared transceivers. This service will be used to interact with the classroom of visually impaired people as well as to enhance users' experience through the sense of touch. This service will use low latency, better than best effort streaming protocols of high priority and feedback control [11, 12].

5. **3-dimensional design and modeling service.** This service will enable the student to have his own virtual workshop where he can use object-oriented toolboxes to design objects. This service will enable the student through a 3d-scanner to transfer via scanning, matter of the real world in the ideal while using 3D printers to implement constructions of the ideal world [18].

6. **Virtual reality recording service.** This LMS service will provide audiovisual recording of the virtual classroom and classroom interactions using 3d user model avatars. Part of the virtual reality recording service is the On-demand playback holographic service of virtual reality context. This LMS service will offer on-demand content and 3d user actions either in the real or virtual world, by showing a past virtual reality recorded session. Part of the recording service, will also be the 3D avatar creation component and the 3D lab presentation module that will be used to project the virtual world to an external user by using projection equipment [18, 19].

7. **Virtual course student assessment service.** This service will use intelligent algorithms developed by this project as well as clustering-classification techniques upon sensory and haptic data of virtual class, or visual information recordings, in order to evaluate the response and overall performance of students in that class. The same service will be responsible for the delivery of course self-evaluations and overall student evaluation reports [13, 14].

Within the framework of the proposed interactive class, the teacher will be able to connect and interact through their students' laptops, tablets, mobile phones and interconnected sensors in an isolated and secure network. Table I summarizes the proposed virtual class services and protocols assigned per each service class.

Additionally, students will also be able to interact with each other in a virtual course and exchange information during that virtual course. The proposed system and implemented protocols will deliver a virtual classroom and real-to-virtual interaction, in which multiple interconnected new technologies will be securely integrated and the services protocols involved as specified per service will provide best effort deliveries, low complexity and high scalability.

SIIC service	Delivery Requirements	Protocols used
Virtual Class service	Virtual class 3D context-real-time Chat services – connectionless over HTTP or UDP LMS asynchronous components – connection oriented over HTTP LMS streams-synchronous components over RTP or HTTP	HTTP – TCP UDP RTP-RTCP NAFCA [4]
Cognitive Service and Augmented sensory services	Asynchronous unreliable Asynchronous reliable Periodic unreliable Periodic reliable	SNMP-UDP CoAP MQTT MESETP [17]
Positioning service	Differentiated- reliable stream, unreliable periodic according to control feedback	UDP RTP MESETP [17]
Haptic service	Little payload, high priority, low jitter, low latency protocol with feedback	A proposition have been made by authors of [11], [12], for a synchronous protocol that modifies its payload and measurements sensitivity according to network conditions

Table 1: Proposed SIIC Services and Protocols Utilised per Service Summary table

V. Implementation Plan

Our research has two key tools for verifying the reliability of the results and the progress of the research project. The first and fundamental tool for verifying the reliability of the results from our research as well as the progress made by the research team will be internal quality control. The internal quality control will be the main responsibility of the research team and will be carried out mainly by the Scientific Director of the Research, as well as by the collaborating institutions that will contribute to the realization of the specific project. The second and "final" tool, as could be described, is a tool for verifying the reliability of the results and the progress of the project is the external quality control, which will be carried out by the collaborators who can access the educational environments and implement "pieces" of the progressive development of the project.

Regarding hardware implementation and computing power along with computational burden, the integration of the aforementioned wireless technologies must be adjusted to constructing an indoor wifi Local Area Network consisted of its essential parts: a wireless network adapter, a wireless router with access points spread across the rooms of the building that facilitates the smart classrooms, proper types of antennas for indoor applications and multiple relays that aim to amplify the propagating signal to reach its destination with sufficient power.

The network adapter aims to improve network performance. The router accompanied by wireless protocols such as 802.11ac is an essential part. Wireless antennas increase the network coverage, while repeaters ensure successful file delivery.

A final remark concerning computational power efficient software solutions must be used together with hardware able to gather, compress and process data with the minimum computation power and burden.

VI. Conclusion and Future Work

The key to achieving this proposal is the ability to develop an evolutionary study, since it will be possible to test what benefits a modern and interactive class can offer, with the specific characteristics and the specific mode of study, in teaching, with the help of specialized associates and team members. Thus, we will be able to contribute in the improvement of the modern educational process, taking into account real scenarios and studies. In this way, we will be given the opportunity to gradually improve the final deliverable as it will be used in real learning conditions. Furthermore, as a case study for the future work of this project will be a smart-interactive laboratory-class and an educational software that can be used in more than one language. The statements above together with the following two paragraphs provide the straightforward benefits of applying the proposed technology.

The proposed SIIC system virtual services and protocols is an authors ongoing research project eligible for funding at ELIDEK (Hellenic Foundation for Research & Innovation) p.III.

This proposal will help pupils to come closer to the scenarios of the modern educational process and to come up with modern supervisory tools. Additionally, because of the "special" nature and laboratory equipment will be given the possibility for remote access to augmented - virtual reality [18-20] by educational institutions all over the world. In this proposed system architecture, there will be a plethora of educational tools/experiments that will use the augmented virtual reality to maximize the learning outcomes. Thus, schools of primary and secondary education will be able to connect to the augmented reality lab via the Internet from their school's computer labs with the help of their teachers, by the use of our propose system architecture.

The social impact of our proposal will be particularly important, as all schools in Primary and Secondary Education will be able to access either physically or online in this experimental laboratory-class. Also, knowledge will not only be acquired through the traditional teacher-centered way of teaching, but emphasis will be placed on experiential learning based on the modern equipment of the laboratory. In an educational learning scenario that we will deal with, a proactive role in experiential learning will be given to the sense of touch, using haptic devices [11], along with the sense of hearing and vision. All students' interaction with the system will be recorded and new teaching methods will emerge through the study of these behaviors.

VII. References

[1] C. Stergiou, K. E. Psannis, “Efficient and Secure Big Data delivery in Cloud Computing”, Springer, Multimedia Tools and Applications, vol. 76, issue: 21, pp. 22803–22822, November 2017.

[2] C. Stergiou, K. E. Psannis, "Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey", Wiley, International Journal of Network Management, pp. 1-12, May 2016.

[3] C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, “Secure integration of IoT and Cloud Computing”, Elsevier, Future Generation Computer Systems, December 2016.

[4] G. Kokkonis, K. E. Psannis, M. Roumeliotis, Y. Ishibashi, “Efficient algorithm for transferring a real-time HEVC stream with haptic data through the internet”, Journal of Real-Time Image Processing,Volume 12, Issue 2, pp 343–355, August 2016.

[5] I. Kakalou I., K. E. Psannis, P. Krawiec, R. Badea, “Cognitive Radio Network and Network Service Chaining towards 5G: challenges and requirements”, IEEE Communications, September 2017.

[6] K. E. Psannis, “Radio Resource Allocation on Complex 4G Wireless Cellular Networks”, 4th International Conference on Mathematical Modeling in Physical Sciences, Session: Statistical Physics and Applications, Mykonos, Greece, June 5-8, 2015 (http://icmsquare.net/).

[7] K. E. Psannis, “Adaptive Layered Segment Algorithm for Media Delivery over 4G LTE Wireless Cellular Networks”, IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB2013), Brunel University, Uxbridge, West London, UK, June 2013.

[8] K. Tsarava, K. Moeller, N. Pinkwart, M. Butz, U. Trautwein, M. Ninaus, “Training computational thinking: Game-based unplugged and plugged-in activities in primary school”, Proceedings of the 11th European Conference on Game Based Learning (pp. 687-695). Reading, UK: Academic Conferences and Publishing International Limited, 2017.

[9] E. Markaki, K. Tsarava, "Collaborative Activities with Cloud Computing Tools for Teacher Training in the Educational Use of Youtube" (2015). 4th Panhellenic

Scientific Conference “Integration and Use of ICT in the Educational Process”, 30th October-1st November 2015, Thessaloniki, Greece.

[10] K. Tsarava, S. T. Halkidis, P. Venardos, G. Stephanides, “ Teaching basic calculus using SAGE”. Strategic Role of Tertiary Education and Technologies for Sustainable Competitive Advantage, 2013.

[11] G. Kokkonis, K. E. Psannis, M. Roumeliotis, S. Kontogiannis, Y. Ishibashi, “Evaluating Transport and Application Layer Protocols for Haptic Applications”, IEEE International Symposium on Haptic Audio-Visual Environments and Games (HAVE2012), Munich, Germany, October 2012.

[12] G. Kokkonis, K. E. Psannis, M. Roumeliotis, S. Kontogiannis, “A Survey of Transport Protocols for Haptic Applications”, 16th Panhellenic Conference on Informatics (PCI 2012), Piraeus, Greece, October 5 - 7, 2012.

[13] I. Kazanidis, S. Valsamidis, S. Kontogiannis, A. Karakos, "Courseware Evaluation Through Content, Usage and Marking Assessment", Research on e-Learning and ICT in Education, Springer ISBN: 978-1-4614-6501-0, Jun 2014, pp. 149-161, 2014.

[14] S. Valsamidis, S. Kontogiannis, I. Kazanidis, A. Karakos, “E-Learning Platform Usage Analysis”, Interdisciplinary Journal of E-Learning and Learning Objects (IJELO), vol. 7, issue 1, ISSN 1436-4522 , Oct. 2011, pp. 185-204.

[15] S. Valsamidis, S. Kontogiannis, I. Kazanidis, T. Theodosiou, A. Karakos, “A Clustering Methodology of Web Log Data for Learning Management Systems”, Journal of Educational Technology and Society (ETS), vol. 15, issue 2, ISSN 1436-4522 , Jul. 2012, pp. 154-167, 2012.

[16] S. Kontogiannis, A. Karakos, "ALBL: An Adaptive Load BaLancing algorithm for distributed web systems", International Journal of Communication Networks and Distributed Systems, vol. 13 issue 2, July 2014, pp. 144-168.

[17] D. Tomtsis, S. Kontogiannis, G. Kokkonis, I. Kazanidis, S. Valsamidis, “Proposed cloud infrastructure of wearable and ubiquitous medical services”, in Proc. Of the 5th Internation conference on Digital Information Processing and Communications (ICDIPC), IEEE proceedings, pp 213-218, 2015.

[18] E. Gounopoulos, S. Kontogiannis, I. Kazanidis, S. Valsamidis, “A framework for the evaluation of multilayer web based learning”, in Proc of 20th Panhelenic conference on Informatics (2016), ACM proceedings, ISBN:978-1-4503-4789-1, pp. 161-164, Nov. 2016.

[19] E. Gounopoulos, S. Kontogiannis, S. Valsamidis, I. Kazanidis, “Blended Learning Evaluation in Higher education courses”, KnowledgeE, vol. 1, No. 1, pp. 385-399, ISSN: 2518-668X, 2017.

[20] I. Kazanidis, S. Valsamidis, S. Kontogiannis, A. Karakos, "Courseware Evaluation Through Content, Usage and Marking Assessment", Research on e-Learning and ICT in Education, Springer ISBN: 978-1-4614-6501-0, Jun 2014, pp. 149-161.