Chapter 1

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

C.L. Stergiou,[1] A.P. Plageras,[1] K.E. Psannis,[1] B.B. Gupta[2]

[1]Department of Applied Informatics, University of Macedonia, Thessaloniki Greece, c.stergiou@uom.edu.gr, a.plageras@uom.edu.gr, kpsannis@uom.edu.gr
[2]National Institute of Technology Kurukshetra, India, bbgupta@ nitkkr.ac.in

Abstract. The 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 Internet of Things (IoT) raises in the sector of networks and telecommunications with specifically 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, with the aim 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, we survey IoT and Cloud Computing technologies with 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.

Keywords. Internet of Things, Cloud Computing, Big Data, Security, Privacy.

1.1 Introduction

"*Internet of Things*" (IoT) is a novel technology which op-erates in the sector of telecommunications. IoT could be defined by many researchers as "the network of devices, vehicles, buildings, and other items which are embedded with sensors, and there are connected to the network, permitting these objects to gather and interchange data" [1] [2] [3]. Over the next years, a flare in the number of connected devices as well as located sites, and the functions they will perform, are expected. Regarding the data used in a wireless network there are security and privacy issues that need to be addressed. The problem with security and data privacy in everyday life could be solved or could be minimized with 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 structured and unstructured form [4] [5]. BD usually uses CC as a base technology in order to operate. Similar to this, another technology that could be used as a base technology is the Edge Computing (EC).

IoT could be settled as a type of network of physical objects or things which are embedded with software, electronics, sensors and connectivity that enables them. Due to that, IoT achieves greater rate and service by transmitting data with operators and various inter-connected devices [6] [7] [8].

An approach has been made by researchers in [9], in order to help other researchers who are interested in se-curity issues. This approach provides an IoT security analysis of the recent security research activity and a novel IoT framework that is validated through a case study. The authors of this paper have shown through their work that the evolution of autonomous objects raises security threats.

Thus, the need of "*cloud*" support has become ineffi-cient due to the intensive computations, the mass storage, and the security issues. Some examples include limited storage capacity, communication capabilities, energy and processing. Inefficiencies like these have motivated us in order to find a model for the combination of CC and IoT. As a "*base*" technology, Cloud Computing consolidates various technologies and applications to get the maxi-mum capacity and performance of the existing infrastruc-ture [10] [11] [12].

On top of that, Mobile Cloud Computing (MCC) made its appearance, as a relative version of Cloud Computing, and it was improved by new developments in the field of "Cloud Computing". The latter aims provide access to data and information from anywhere at any time by obliterat-ing the need for hardware equipment [2] [13] [14] [15]. More specifically,

MCC is defined as an integration of cloud and mobile computing rendering mobile devices more resourceful. It is also a contemporary approach to innovative services for firms and institutions. CC can be used as a useful base for both Internet of Things and Video Surveillance technologies and provide improvements on their function [16] [17] [18].

Moreover, Cloud Computing aims to offer access to in-formation and data from anywhere at any time, without the restrictions of the need for hardware equipment [11] [19] [20] [21]. As a result of the operations of CC, it could be used as a base technology for IoT and for several technologies in the telecommunications field, and could also provide improvements on their functions.

In addition to this, CC additionally used to be a base technology for other technologies due to its types of services [11] [21] [22]. One of those is the Big Data. BD is a term used to describe the expected, due to the connected to the Internet devices, rapid increase in the volume of data production. Subsequently, these large amounts of data could be defined as "a broad term for data sets so large or complex that traditional data processing applications are inadequate" [12] [21]. Furthermore, BD is often associated to the use of predictive analytics or certain advanced methods to extract knowledge from the data. Rarely, are also related to a particular size of set of data [4] [5]. Precision in BD could result in more confident decision making, and better decisions may drive in increased operational efficiency, reduced costs, and minimized risk [4]. From this scope, it can be observed that BD is now equally important both for business and internet. This happens because more information drives to more accurate analysis [11]. The real problem is not that the large amounts of data have been obtained, but whether they have any value or not. Hopefully, by predicting that organizations would be able to acquire information from any source, harness the relevant data, and analyze them in a specific way in order to get quick answers, the following should be achieved: 1) reduce costs, 2) reduce time, 3) produce new items and optimize their offerings, and 4) take more ingenious decisions [7].

Last but not least, since we are talking about BD, IoT, and CC/MCC many researchers tried to figure out ways for securing these sensitive/personal data. The security problems still remain a challenge since the new technologies are multiplied. Due to this, a security scheme for safe sensitive data transmission over the CC and the IoT de-vices has been proposed in [23]. Specifically, an alternative of RSA (Rivest-Shamir-Adleman) security has been deployed, namely MEMK ("*Memory Efficient Multi Key*") generation scheme, in order to provide support to the data transmitted from the IoT devices to the Cloud and back. This scheme has

been also used by the authors of this paper [23] to boost the efficiency of the memory.

The rest of the paper is divided in sections as follows. Initially, in Section II has been presented a literature re-view related to the conjunction of the technologies mentioned in the introduction section (Section I) of the paper. Subsequently, in Section III there is an illustration of is-sues related with BD and their privacy. In Section IV has been discussed in detail the field of IoT and some of its major functions. Moreover, in Section V the CC technolo-y and its basic characteristics have been presented and analyzed. Section VI illustrates the integration of IoT and CC, and surveys some of the benefits of their integration. Finally, Section VII provides the conclusions of the cur-rent paper, and offers new possibilities for the development of future work.

1.2 Literature Review

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 dis-

cover 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 holis-

tic 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 aggrega-

tion 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*".

1.3 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*".

1.3.1 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].

1.3.2 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 un-known 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.

1.3.3 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].

Table 1.1. Encryption rates of popular algorithms

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

Regarding the Table 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.

1.3.4 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 addi-

tionally 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.

1.4 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).

1.4.1 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.

1.4.2 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 com-

puting 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].

1.4.3 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].

Amplify-and-forward (AF)

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.

1.5 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].

1.5.1 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].

1.5.2 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.

1.5.3 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].

1.6 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 2 presents the features of CC regarding the accessibility of this technology provides when combined with the characteristics of IoT [82] [83].

Table 1.2. Contributions of Cloud Computing in Internet of Things

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 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 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 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.

1.6.1 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].

Table 1.3. Affects of IoT & Cloud Computing security challenges

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

Table 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 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).

1.6.2 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].

Table 1.4. Correlation of BD and CC characteristics

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 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 4 is that the feature *Applications over Internet* contributed from all the key features of BD.

1.6.3 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.

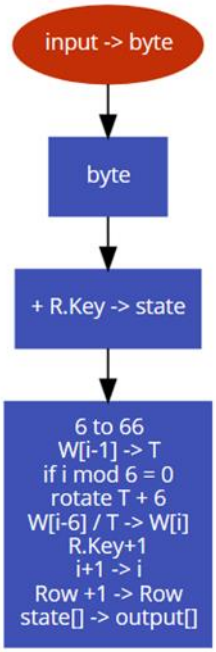

Fig. 1.1. Flowchart of the proposed the procedure implementation

Figure 1 shows the proposed pseudocode representation through a flowchart.

1.6.4 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.

Table 1.5 AES contribution in IoT and Cloud Computing

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**

The Tables 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 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.

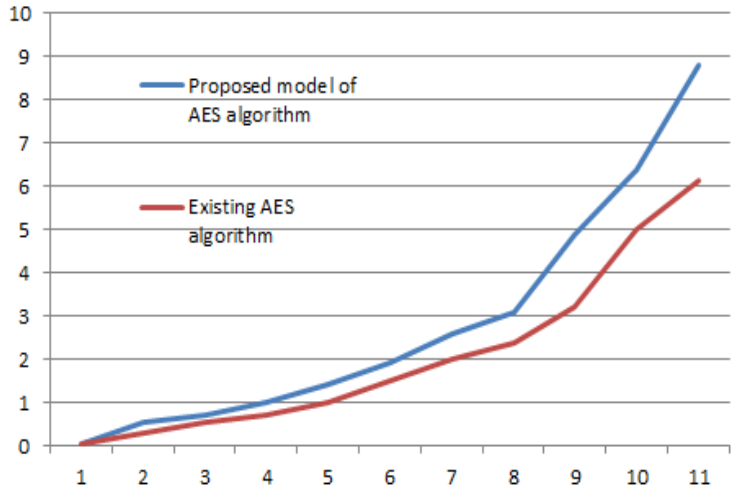

Fig. 1.2. Security level of encryption algorithms of measurement used for the study of AES model algorithm

Figure 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.

Table 1.6. Correlation of BD characteristics with IoT & CC integration model

Big Data Features	**Volume**	**Velocity**	**Variety**	**Veracity**	**Value**
IoT & CC integration model	**X**	**X**	**X**	**X**	**X**

The Tables 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 6 presents that all the characteristics of BD contributed by the integration model of IoT and CC technologies.

1.7 Conclusions

The CC technology provides a number of possibilities, but additionally places several limitations as well. Cloud Computing refers to an infrastruc-

ture 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.

Acknowledgement. The authors would like to thank the anonymous reviewers for their valuable comments and feedback which was extremely helpful in improving the quality of the paper.

References

1. L. Atzori, A. Iera, G. 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]
2. C. Stergiou, K. E. Psannis, B.-G. Kim, B. Gupta, "Secure inte-gration of IoT and Cloud Computing", Elsevier, Future Gener-ation Computer Systems, December 2016.
3. A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, B. B. Gup-ta, "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.
6. J. Mongay Batalla, P. Krawiec, “Conception of ID layer per-formance at the network level for Internet of Things”, Springer Journal Personal and Ubiquitous Computing, Vol.18, Issue 2, pp. 465-480, 2014.
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. Ishi-bashi, “Architecture for Security in IoT Environments”, in Pro-ceedings of 26th IEEE International Symposium on Industrial Electronics, 19-21 June 2017, Edinburgh, Scotland, UK.
9. A. R. Sfar, E. Natalizio, Y. Challal, Z. Chtourou, “A Roadmap for Security Challenges in Internet of Things”, Elsevier, Digital Communications and Networks (DCN), vol. 4, issue: 2, pp. 18-137, April 2018.
10. Y. Kryftis, G. Mastorakis, C. Mavromoustakis, J. Mongay Ba-talla, E. Pallis and G. Kormentzas, “Efficient Entertainment Services Provision over a Novel Network Architecture”. To be published in IEEE Wireless Communications Magazine, 2016.
11. C. Stergiou, K. E. Psannis, "Recent advances delivered by Mo-bile Cloud Computing and Internet of Things for Big Data ap-plications: 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.
13. C. Stergiou, K. E. Psannis, A. P. Plageras, G. Kokkonis, Y. Ishibashi, “Architecture for security monitoring in IoT envi-ronments”, in Proceedings of IEEE 26th International Sym-posium on Industrial Electronics, Edinburgh, Scotland, UK, June 2017.
14. A. P. Plageras, K. E. Psannis, Y. Ishibashi, B.-G. Kim, "IoT-based Surveillance System for Ubiquitous Healthcare," In-dustrial Electronics Society, in Proceedings of IEEE/IECON 2016 - 42nd Annual Conference of the IEEE, October 2016.
15. J. Li, L. Huang, Y. Zhou, S. He, Z. Ming ,"Computation parti-tioning for mobile cloud computing in big data environment", IEEE Transactions on Industrial Informatics, Vol. 11 January 2017.
16. R. Yu, X. Huang, J. Kang, J. Ding, S. Maharjan, S. Gjessing, Y. Zhang, "Cooperative Resource Management in Cloud-Enabled Vehicular Networks", IEEE Transactions on Industrial Informatics, volume: 62, Issue: 12, pp 7938 - 7951, September 2015.
17. D. Agrawal, B. B. Gupta, S. Yamaguchi, K. E. Psannis, "Recent Advances in Mobile Cloud Computing", Wireless Communications and Mobile Computing, December 2017.

18. A. M. M. Ali, N. M. Ahmad, A. H. M. Amin, “Cloudlet-based cyber foraging framework for distributed video surveillance provisioning”, Information and Communication Technologies (WICT), 2014 Fourth World Congress on, Bandar Hilir, Malaysia, December 2014.
19. M. R. Rahimi et al, “Mobile Cloud Computing: A survey, State of Art and Future Directions”, Mobile Networks and Applica-tions, Volume 19, Issue 2, pp. 133-143, March 2014.
20. 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" 46th Hawaii International Conference on System Sciences, pp. 1025-1034, October 2013.
21. C. Stergiou, K. E. Psannis, “Algorithms for Big Data in Ad-vanced Communication Systems and Cloud Computing”, in Proceedings of 19th IEEE Conference on Business Informatics 2017 (CBI2017), Doctoral Consortium, 24-26 July 2017, Thessa-loniki, Greece.
22. A. P. Plageras, C. Stergiou, K. E. Psannis, Byung-Gyu Kim, Brij Gupta, Y. Ishibashi, “Solutions for Inter-connectivity and Secu-rity in a Smart Hospital Building”, in Proceedings of 15th IEEE International Conference on Industrial Informatics (INDIN 2017), 24-26 July 2017, Emden, Germany.
23. C. Thirumalai, H. Kar, “Memory Efficient Multi Key (MEMK) Generation Scheme for Secure Transportation of Sensitive Data over Cloud and IoT Devices”, IEEE, in Proceedings of 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), 21-22 April 2018, Vellore, india.
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. Gup-ta, “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 Man-agement 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 Man-agement System 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 Compu-ting: 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 Com-munications 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. [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 Monitoring 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””, Future 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", Elsevier, Science Bulletin, vol. 64, issue 14, pp. 1025-1029, July 2017.
52. G. Bello-Orgaz, J. J. Jung, D. Camacho, "Social big data: Recent achievements 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 Management, 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 Im-age 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 Mul-timedia 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 Sys-tems, 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, No-vember 2015.

72. Y. Kumar, R. Munjal, H.Sharma,"Comparison of Symmetric and Asymmetric Cryptography with Existing Vulnerabilities and Countermeasures" IJCSMS International Journal of Com-puter Science and Management Studies, Vol. 11, Issue 03, Oc-tober 2011.
73. R. Kaur, S. Kinger, "Analysis of Security Algorithms in Cloud Computing," International Journal of Application or Innova-tion in Engineering & Management (IJAIEM), vol. 3, no. 3, pp. 171-176, March 2014.
74. G. Singh, S. Kinger, "Integrating AES, DES, and 3-DES Encryp-tion 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 Secu-rity 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 envi-ronment", 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 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.
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 applica-tions”, 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 moni-toring: A survey”, Computer Networks, vol. 57, issue: 9, pp. 2093–2115, 2013.