Vaal University of Technology

Your world to a better future

PREDICTION OF TRAFFIC FLOW IN CLOUD COMPUTING AT A SERVICE PROVIDER

By

PRINCE SEKWATLAKWATLA

Student no: 213087766

Submitted in accordance with the requirements for the degree of

MAGISTER TECHNOLOGIAE

IN THE DEPARTMENT

OF

INFORMATION TECHNOLOGY

at the

VAAL UNIVERSITY OF TECHNOLOGY

Supervisor: Prof. Tranos Zuva

Co-Supervisor: Mr Raoul Kwuimi

30 November 2018

DECLARATION BY CANDIDATE

"This is to announce that this thesis is composed by Sello Prince Sekwatlakwatla and has not been submitted to any other higher learning institution but Vaal University of Technology, in partial fulfilment of the M-Tech degree in Information Technology in the Faculty of Applied and Computer Science. I decidedly pronounced that all sources referred to are acknowledged by methods of a comprehensive rundown of references."

Full Names: Sello Prince Sekwatlakwatla

Date: ---------------------------------

Signature: ---------------------------------

DEDICATION

This thesis is committed to GOD Almighty the author and the finisher of my faith. I thank GOD for his guidance, help and protection over me. The thesis is likewise, dedicated to my family and friends.

ACKNOWLEDGEMENTS

To begin with I might want to guide my legitimate gratefulness and thankfulness to my supervisors, Prof T Zuva and Mr. Raoul Kwuimi for their kind words, direction, tolerance, and support during my MTech journey. Their commitment to my work is immense and can't be measured against any standard.

Very special thanks go to my supervisors, without whose motivation and reassurance I would not have engaged in a graduate career in Cloud computing research. Prof T Zuva and Mr. Raoul Kwuimi are parents who made a huge difference in my life. It was under their coaching that I advanced a focus and became interested in research influences. They supplied me with direction and technical support and became more mentor and teacher, than professor.

It was however their, assurance, comprehension and empathy that encouraged me to complete my Master's degree and stimulated me to keep contemplating. Much thanks for liberally and compassionately imparting to me your enormous knowledge in the area of Cloud computing. This empowered me to comprehend and gain knowledge on the subject of research. May GOD BLESS YOU abundantly! I couldn't have envisioned having better supervisors for my MTech studies. I don’t know if I will ever have the capacity to convey my appreciation completely, yet I owe him my endless gratefulness.

Finally, I would like to thank all those people who availed their precious time during the data gathering process of this research report, without them this research wouldn’t be a success.

CONTENTS **PAGES**

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Cloud computing provides improved and simplified IT management and maintenance capabilities through central administration of resources. Companies of all shapes and sizes are adapting to this new technology. Although cloud computing is an attractive concept to the business community, it still has some challenges such as traffic management and traffic prediction that need to be addressed. Most cloud service providers experience traffic congestion.

In the absence of effective tools for cloud computing traffic prediction, the allocation of resources to clients will be ineffective thus driving away cloud computing users. This research intends to mitigate the effect of traffic congestion on provision of cloud service by proposing a proactive traffic prediction model that would play an effective role in congestion control and estimation of accurate future resource demand. This will enhance the accuracy of traffic flow prediction in cloud computing by service providers. This research will evaluate to determine the performance between Auto-regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) as prediction tools for cloud computing traffic.

These two techniques were tested by using simulation to predict traffic flow per month and per year. The dataset was downloaded data from a public website taken from the Center for Applied Internet Data Analysis (CAIDA) database. The two algorithms Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) where implemented and tested separately. Experimental results were generated and analyzed to test the effectiveness of the traffic prediction algorithms. Finally, the findings indicated that ARIMA can have 98 % accurate prediction results while ANN produced 89 % accurate prediction results. It was also observed that both models perform better on monthly data as compared to yearly data. This study recommends ARIMA algorithm for data flow prediction in private cloud computing.

CHAPTER 1

1. Introduction

This thesis reports on traffic flow prediction in cloud computing, according to the need of cloud users for a service provider that meets service level agreements. Over the last few years, the telecommunications industry has seen cloud implementation evolve from a developing technology to an established networking solution that has gained widespread acceptance and deployment. With the quick growth of cloud computing technology, network traffic is due to increase and thus experience network congestion. In this regard, the network administrators have to keep abreast with the latest techniques that will help them maintaining the operational status quo of the organization and also running the network at optimal levels (Dinh et al., 2013).

The benefit of traffic prediction and analysis of data in cloud computing includes quicker execution, continuous web presence, more collaboration, lower subscription cost and load balancing with minimum infrastructure (Dinh et al., 2013). Therefore, traffic management and analysis issues should not be ignored. These benefits are perfect inspirations for organizations to transfer their data to the cloud and take into consideration the possibility of predicting the future to be able to manage and provide constant resource to their clients. However, the investigation of the data traffic in cloud computing has by many trials. These trials stem from the difficulties in calculating, assessing and predicting the traffic to the data centers, particularly when they are composed of several hundreds of servers (Zander and Mahonen, 2013).

It is conjectured that traffic flow prediction will become a dominating trend for required cloud services in the future of cloud data-centers. This research tackles the emerging research theme of providing advanced monitoring functionalities for cloud services to help service providers and clients to manage Cloud and future distribution of resources.

The intention of this research is to investigate the algorithm used to accurately predict traffic flow in the cloud. This will ensure that there is efficient and effective allocation of resources to enable the client to get the best service from cloud computing. To achieve this the Auto-Regressive Integrated Moving Average (ARIMA), and Artificial Neural Networks (ANN), is proposed as prediction tools for cloud computing traffic flow. This chapter introduces the context in which traffic prediction in cloud computing models are used, and presents a problem statement, research questions and the summary of the chapter.

1.1. Problem Statement

Cloud computing provides improved and simplified IT (Information Technology) management and maintenance capabilities through central administration of resources (Prince et al., 2016), and various companies are adopting this new technology. In the absence of an effective prediction tool of cloud computing traffic, allocation of resources to clients will be ineffective, thus driving away cloud computing users.

In this regard, the network administrators have to keep abreast with the latest techniques that will help them maintain the operational status quo of the organization and also run the network at optimal levels (Santos et al., 2009; Sadeghi et al., 2010; Grobauer et al., 2011). System administrators must administer control and association strategies to enhance the system in an exceptional way, so as to meet the requirements of clients. Arranging the activity stream is critical in system administration to ensure that the system is online.

In cloud computing the inconsistent network traffic flow leads to difficulties of predicting the network resources that are appropriate to service the needs of all network clients at that point in time. Service providers are unable to allocate the resources to cater for the needs of clients (Vlahogianni et al., 2014). Inconsistent traffic flow leads to clients complaining about slow system times, application timeout and high usage of bandwidth during peak times.

The Current issues are extraordinary system frameworks. These frameworks move large volumes of information inside the cloud and they can cause system bottlenecks as opposed to CPU-or memory-bottlenecks. Examples of these extraordinary frameworks are Hadoop occupations (Radhakrishnan et al., 2012), workloads, reinforcement administrations (Manvi and Shyam, 2014) and logical or scientific calculations. Such PC programs work crosswise over many machines (Lagar-Cavilla et al., 2009) and are on both open cloud and business/extend datacenters.

1.2. Research Question

Cloud computing environments raise certain concerns within the network management area. In this regard, the main research question for this study is: How can cloud computing usage be predicated accurately?

It is envisaged that the answers to the following sub-questions will aid in answering the main research question.

1 How is cloud computing data flow prediction measured? (in literature)
2 What techniques can be used to predict accurate traffic flow in cloud computing?
3 How can the effectiveness of the technique be measured?
4 Which technique can be recommended to service providers as being more efficient?

1.3. Research Objectives

The key objectives, crucial for this research, were identified as follows:

1. To investigate traffic flow prediction techniques of cloud computing in literature.
2. To propose effective traffic flow prediction techniques(s) for cloud computing service providers

3. To measure the effectiveness of the proposed technique.
4. To perform traffic flow prediction in cloud computing using the proposed technique

1.4. Research Contributions

The contribution of this research, are as follows.

1. An appropriate prediction algorithm is recommended to be used by cloud computing services providers in order to avail the resources to their clients.
2. Deploying this technique to companies and industries will enhance system productivity and future prediction of resources.
3. Research papers were written and published

1.5. Thesis Outline

This thesis is arranged as follows:

Chapter 1 Introduction.

Chapter 2 Reviews of related Research on Traffic prediction and a discussion based on theoretical framework for this research work.

Chapter 3 Focuses on methodology, data collection and system implementation.

Chapter 4 Focuses on design, experiment results and discussion based on Traffic prediction System.

Chapter 5 Conclusion, contribution and future work. The achievements, shortfalls and future endeavours are discussed.

1.6. Summary of chapter 1

This chapter has explained in detail the importance of traffic flow in cloud computing and introduced the importance of traffic flow prediction in cloud computing. The chapter gives an overview of the study which includes, the problem statement, research question, research goals and the thesis outline. The following chapter focuses on the literature review.

CHAPTER 2

2. Literature Review

Cloud computing is a complete and full distributed system that allows on-demand network access to computing resources with minimal management effort or service provider interaction in which a group of IT companies are virtualized, managed or used to calculate power, storage and facilities. Facilities are delivered on demand to external customers over the internet (Grobauer et al., 2011; Sadeghi et al., 2010;Santos et al., 2009). There is a steady increase in interest in cloud computing, especially to own a cloud (Mell and Grance, 2011). This is affected by the following factors: reduction in hardware cost, growth in computing power and storage capacity and the exponentially rising in data size in technical instrumentation and simulation. Simulation and software development in public clouds are more expensive than those in private clouds (Grossman and White, 2012).

Cloud computing is a utility computing model, where customers using the service, pay on the basis of usage and pricing is determined by the service provider. Virtual machines are run on physical machines that consume energy and other resources. The customer has to pay a fee to support the cost of running the service, the cost of employees who monitor the machines and the cost to replace broken down hardware. Providers can charge users by the hour, they use virtual machines to measure the traffic that has passed and the data storage cost.

There are three types of administration provided by various cloud specialist organizations, namely, i IaaS - Infrastructure as a Service, ii Saas - Software as a Service, and iii PaaS - Platform as a Service. Amazon EC2 provides the IaaS administration and this postulation concentrates on building structure for Infrastructure as a Service. It gives cloud computing innovation client's free rein to have their own working framework with required programming (Grossman and White, 2012).

Cloud computing gives free rein for changing the virtual machine and offering potential outcomes to interface with the virtual machine. The most imperative concerns in cloud computing is the technical issues where information and data on the cloud can be accessed any time and from anywhere as there are times when the system can serious malfunction. Businesses should be aware of the fact that this technology is always prone to outages and other technical issues. Even the best cloud service providers run into this kind of trouble, in spite of keeping up high levels of maintenance (Mell and Grance, 2011). The favorable factors for the expansion in the request of cloud computing appears in Figure 2-1, and includes savings, efficiency, accessibility, flexibility, innovation and opportunity.

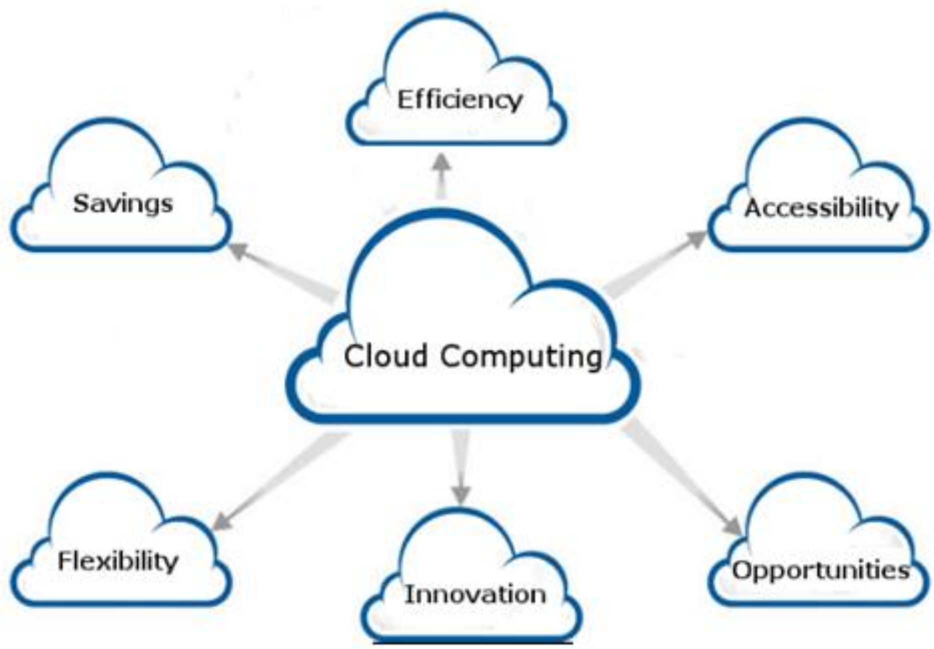

FIGURE 2-1: cloud computing benefits (Buyya et al., 2010)

In cloud computing the issue of conflicting network traffic flow prompts challenges of anticipating the system assets that are suitable to benefit the customers by meeting a Service Level Agreements (SLA). Besides, there is a need for traffic expectations with a specific end goal, to alter the current explanatory models to anticipate the volume of future and current traffic (Gill et al., 2011).

The benefit of assessing traffic flow issues in cloud computing will assist service providers to manage their system, for example, overall internet and its forthcoming frameworks such as cloud computing. Furthermore, there is a need to highlight numerous feedback methods and determine which is most appropriate to address the problem of inconsistent network traffic flow. In cloud computing performance assessment schemes can be used for accurate traffic prediction systems. The following section focuses on cloud computing service.

2.1. Cloud computing Services

Cloud services are various specialist organizations offering the opportunity to utilize cloud computing to settle computationally concentrated errands and giving practically infinitive measure of assets to utilize in order to manage organisations. There are a few deliberation layers of administrations given by these providers as shown in Figure 2-2: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) (Mell and Grance, 2011).

FIGURE 2-2: services of cloud computing (Mell and Grance, 2011)

Figure 2-2 shows three cloud computing services which are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS)

As indicated by Mell and Grance (2011), software as an service is a product circulation show in which programming and its related information are facilitated midway, typically in the cloud and are available to clients by utilizing a thin client. SaaS has become a common delivery model for most business applications. These applications can be obtained by paying a fee. In 2010, SaaS sales reached 10 billion US dollars (Mell and Grance, 2011). SaaS software can be used by a large number of customers with little resources making it possible to run it at a low price.

These services are available for a membership charge either on a monthly or a yearly basis. In case of an expansive client base, some of the suppliers can offer applications with a free premium model. This implies that a portion of the administration is freely available. Some different SaaS suppliers like Google can provide applications to clients, since they derive an income for exchange sources, such as advertising. Promoting is one regular approach to provide free fundamental programming for the end users.

Typical SaaS applications provided by Google are Gmail (an e-mail application), and Picasa (a picture application). The challenge that service providers face is the increasing data traffic in cloud computing.

2.2. Benefit of Software as a Service

The following are some of the benefits of Software as a service (Mell and Grance, 2011):

- No early setup costs - PC projects are set up to be utilized when the client subscribes;
- Pay just what you use - if programming is requested for a restricted period then it is paid for over that period and memberships can as a rule be ended at any time;
- Stockpiling or additional - if a Client think that they need more they can contact these on welcome without expecting to put in new programming or equipment;
- Gadget similarity - SaaS PC projects can be recovered through any gadget, which makes it ideal for the individuals who utilize various gadgets, such as, web enabled telephones and tablets. Individuals who do not have these gadgets, persistently make use of a portable PC;
- Available from anywhere - as opposed to being constrained to associations on isolated PCs, a PC program can be recovered from anywhere with a web enabled gadget; and
- Customisation is available - it can be changed to suit the required needs and calling/marking of an obviously expressed/specific client.

2.3. Examples of Software as a Service

In Figure 2-2 below are examples of the services that you can use on a daily basis from SaaS:

FIGURE 2-3: services of cloud computing(Mell and Grance, 2011)

2.4. Tools of Software as a Service

Programming expectations with Self-versatile Arrangement (SPSA) determine accurate predictions and groups the clients programming requirements (Herbst et al., 2014). There are many types of hardware that can be utilized as a part of software as an administration. Figure 2-4 lists some of devices that can possibly be utilized to actualize a product as an administration, some of these are free while others are pay-as-you-utilize.

FIGURE 2-4: Tools of Software as a Service (Herbst et al., 2014)

- WordPress (site) - Create a premium blog or website with moderate subjects, free modules, and a large amount of engineer support;

- Tumbler (blog) - enables individuals to deliver instance, music;

- Google Analytics (web examination) – users are able to track video and person to person communication;

- Feedly (RSS peruser) - numerous plan choices and versatile applications;

- Trackur (online networking observing) - is an operational standing administration and web-based social networking checking gadget;

- Hootsuite (online networking engagement) - makes it easy to engage with individuals over the greater part of your web-based social networking accounts proficiently and reliably, over your association;

- Zendesk (help work area) - Customer administration to help client to rest or open records;

- Kashoo (bookkeeping) - bookkeeping programming for business;
- Basecamp (extend administration) - is a venture administration device that can be utilized by many web activities;
- PayPal/Braintree (installment preparing) - Quick and Easy Payment Processing;
- Fresh books (invoicing) - the best invoicing and charging programming intended for entrepreneurs;
- Security (malware security) - a comprehensively archived expert in all matters identified with site security, with specialization in WordPress Security;
- Iubenda (security strategies)- The most exquisite and compelling approach to produce a protection arrangement;
- Hostdime (facilitating) - server facilitating instrument, affiliate facilitating;
- Hivelocity (facilitating two) - custom server administration instrument;
- Vimeo (video player) - The player defaults to HTML5 to provide the most exquisite review understanding;
- Pingdom (site uptime) - makes web execution administration simple. Screen your destinations uptime, execution, exchanges and others;
- MaxCDN (content conveyance organize) - a substance appropriation arrange (CDN) supplier;
- Envato (web inventive) - creative website template for law, lawyers, attorneys and legal agencies; and

- iStockPhoto - royalty free stock photos, vector art photographs, stock footage and audio for print and use on presentations.

As much as there are devices in SaaS, there is an increase of traffic each day creating traffic flow congestion. LaCurts, (2014) proposed that with Cicada's workload application the aim was to keep in mind the end goal and to help suppliers and clients to enhance and figure out the application's system needs.

2.5. Platform as a Service

This classification of cloud computing gives an assurance and indication of what administration is. In the layered model of cloud computing, the PaaS layer lies between the SaaS and the IaaS layers and it offers assistance between these two frameworks without the additional cost of purchasing other equipment.

PaaS officers encourage and create frameworks without the cost and complexity of purchasing and dealing with the concealed equipment.

There is an assortment of combinations of administrations available to bolster the application improvement life-cycle (Buyya et al., 2010). A portion of the suppliers offer administrations where clients can choose any database, any programming dialect and any working framework to advance productivity. One common example of PaaS is Google Application Engine that permits creating and facilitating web PC programs in Google-overseen datacenters.

These applications are effectively versatile. As indicated by (Buyya et al., 2010) the Google Application Engine is free up to a specific level of usage. Extra fees are charged for additional capacity, transfer speed or, occasionally, extra hours required by the application. The Google Application Engine is free and helpful for important applications. More charges are incurred after some time for additional capacity, radio capacity or extra hours requested by the online framework, where additional work, cash, confirmation, etc. is required.

2.6. Benefits of Platform as a Service

- Flexibility - features can be changed if required;
- It is not important to deal with the synopsis of new declarations of the advancement or essential programming;
- It is not important to office servers. The PaaS cloud provider will deal with that;
- It is not important to oversee reinforcements. The PaaS cloud provider will deal with that; and
- It is not important to deal with the essential server farm. The PaaS cloud provider will deal with that.

2.7. Examples of Platform as a Service

There are many companies that offer a platform as a service, some of the tools of platform as a service are shown in Figure 2-5.

FIGURE 2-5: Tools of Platform as a Service (Buyya et al., 2010).

- Cloud Foundry - introduced by VMware, delivers diverse products and services as a platform as a service;

- Cloudify - mainly dedicated on automation, but also focuses on deployment, monitoring, auto-scaling of application stacks based on usage and remediation;

- OpenShift – developed by Red Hat. The company presently supports a private cloud version of the software;

- Stackato - it deals with customizable application stores; and

- WSO2 Stratus - it provisions more essential services than the other accessible PaaS options today and is a good option for enterprises.

2.8. Infrastructure as a Service

As indicated by (Buyya et al., 2010) IaaS gives cloud suppliers the opportunity to use virtual machines, that can utilize any working framework they like and introduce foundations without any preparation, with fundamental programming. A popular IaaS provider is Amazon EC2 where infrastructure building and validation is done in an IaaS cloud computing model. IaaS bill users by the hourfor every instance they uses each virtual machine, bandwidth, storage and other services, as soon as they are activated (Buyya et al., 2010).

There are many other services that can accompany existing infrastructure. Users can use elastic load balancing service with auto scaling as all the necessary services are there to support scalability of the application on the cloud. Some of the services are provided for free, e.g. gathering CPU. Memory and network usage is provided without any additional charge, but the traffic flow values are collected within 5 minute intervals. If customers want to measure the performance frequently or use custom metrics, they have to pay for the extra service. To advertise the service and attract new customers, Amazon EC2 provides new users free access for hours per month for one year.

2.9. Benefit of Infrastructure as a Service

- Flexibility and scalability - businesses can create a cloud service and develop a system that meets their needs initially. IaaS permits businesses to choose services that supports the commercial use and delivers the greatest support for clients;

- Innovative applications and services - businesses must use the latest technologies and services without needing to spend more funds to market their product;

- Companies can develop businesses that apply IaaS and have the single ability to subcontract to another company that can handle their organization's needs; and

- Affordability - organizations can use IaaS framework without outsourcing any other organisation.

2.10. Examples of Infrastructure as a Service

Cloud computing is a preparatory model that empowers both data innovation foundation and programming to be dispersed transparently over the Internet as an administration. Most of the IT companies or researchers have designed algorithms to predict traffic accurately. The problem to cater for the needs of clients still persist due to the lack of accurate predictions (Wan et al., 2014). Figure 2-6 demostrates an example of infrastructure as a service.

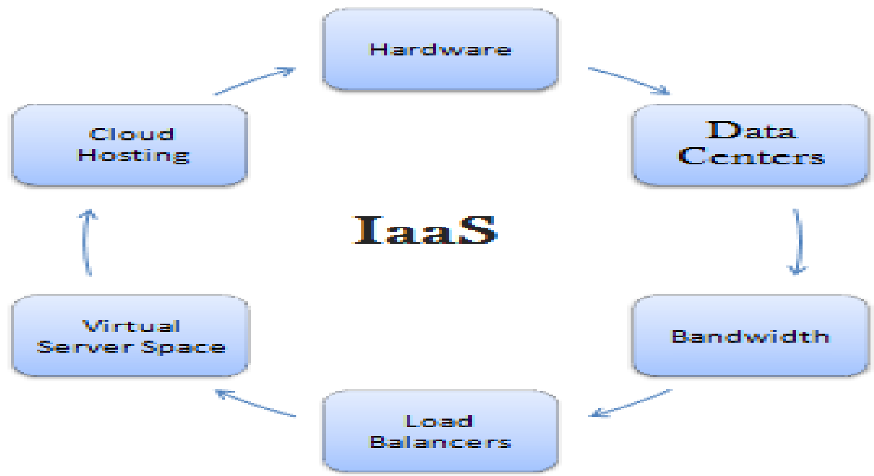

FIGURE 2-6: Infrastructures as a Service

2.11. Types of cloud computing

Cloud computing offers planners and IT offices with the capacity to concentrate on what materials are essential and prevent undistinguishable work such as buy, upkeep, and capacity. Cloud administration has the potential for arrangements to give the client different levels of control, adaptability, and administration. The cloud computing administration is partitioned into three parts, namely, private, public and hybrid cloud which is a blend of two cloud administrations (see Figure 2-7 below).

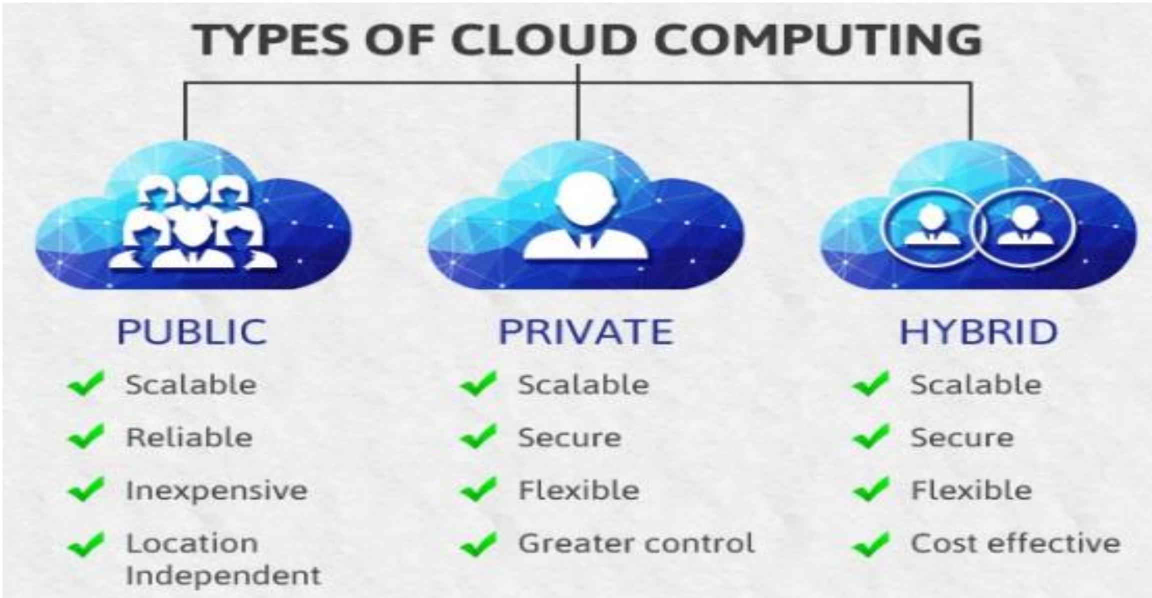

FIGURE 2-7: Types of cloud computing (Mell and Grance, 2011)

2.12. Public Cloud

This is a specialist provider that provides helpful services, for example, it provides PC projects and capacity and is accessible to everyone over the Internet. Open cloud service may be free or offered on the basis of getting payment with every utilization. Examples of open cloud include Amazon Elastic Compute Cloud (EC2), IBM's Blue Cloud, Sun Cloud, Google Application motor and Windows Azure Services Platform (Mell and Grance, 2011). For clients, these types of clouds will provide the most affordable method to set-up in light of the fact that equipment, PC programs and radio expenses are paid by the supplier. It is an advantage for every cloud user to show that the available capacity is utilized. However there are constraints, such as, service level agreement specificity, erroneous expectation of traffic, mistaken allotment of resources, security and other constraints that must be addressed in this cloud.

2.13. Advantages of public cloud computing

- It offers better scalability;

- It is cost efficient and helps save money;
- It offers consistency in that a single point of failure will interrupt your service;
- Services like SaaS, PaaS, and IaaS are available on public cloud platforms and can be accessed from anywhere through any internet service provider; and
- It is location independent – the services are available anywhere the client is located.

2.14. Disadvantage of public cloud computing

- No stringent protocols regarding data management;
- Privacy or security cannot be controlled;
- Complex applications are not accessible; and
- Unavailable adaptability as the stage is determined solely by stage supplier.

2.15. Private clouds

Private clouds are a model of cloud computing where information technology services are provisioned over local IT infrastructure for the dedicated use of a single organization. The objective of a private cloud is not to offer it as an administration to clients, but rather to utilise advantages of the cloud without surrendering the control of maintaining your own server farm.

Private cloud can be costly and is generally impossible for the normal Small-to-Medium established business and is therefore mainly used by large organizations. The main concern of private clouds resolve around security and maintaining important things inside the firewall (Mell and Grance, 2011).

2.16. Advantages of private cloud computing

- Offers greater security and privacy;
- Offers more control over system configuration as per the company's needs;
- Greater reliability when it comes to performance;
- Enhances the quality of service offered to the clients; and
- Saves money to the organisations.

2.17. Disadvantages of Private cloud

- Expensive when compared to public cloud; and
- Requires IT expertise in the organisations.

2.18. Hybrid clouds

Hybrid techniques are partnerships that can control and oversee private clouds while relying upon people in general cloud as and when needed, such as pinnacle stages singular PC projects, or parts of PC projects can be moved to the public cloud (Islam et al., 2012).

The capacity to keep up and off disaster improvement site for most associations is incredible expensive,. There are lower hybrid clouds which takes a toll on arrangements and on decisions to drop down the range that an association gets, and therefore the capacity to recoup information rapidly reduces (Islam et al., 2012). Cloud based Disaster Recovery (DR)/Business(BC) (continuous, steady quality) administrations enable associations to contract failover out to a managed services provider that keeps up multi-occupant for DR/BC, and has practical experience in getting business back online quickly (Islam et al., 2012).

2.19. Advantages of hybrid cloud computing

- It is scalable - ability of a process, network, software or organization to grow and manage increased demand;

- It is cost efficient - organization can save money and performing activities in a better way;

- Offers better security - advanced security of data in an organisation; and

- Offers greater flexibility - the ability to move joints effectively through a complete management of data in organisation.

2.20. Disadvantages of Hybrid Cloud computing

- Infrastructure dependency - highly interconnected and mutually dependent; and

- Possibility of security breach through public cloud.

Data in cloud computing is most affected by system transmission capacity, reaction time, availability of a service, performance unpredictability, lease postponement in information exchange (Hu et al., 2010). In cloud computing the traffic forecast assumes an imperative part in predicting the system assets that is suitable to benefit the necessities of system customers. Zhang et al., (2010) concentrated on asset prerequisite expectation in appropriated frameworks like lattices and mists. In cloud computing, the execution information for each occupation running in was put away in an execution history. Recently submitted engagements were then precisely examined to discover its duplicates from the execution history and in view of the information, put away in the execution history; the valuable supply required thing of the new occupation was anticipated.

This study investigated the diverse operator established arrangements regarding traffic forecast in cloud computing. This research was based on the necessities for the traffic forecast in the cloud computing stage.

2.21. Traffic prediction in cloud computing

In order to help specialist organizations and their customers to predict an application's system traffic in cloud computing, the service provider needs to make exact estimates for both normal requests and the aggregate sum of information (Hu et al., 2010). An application will require to move over in future and most of the information and application should move over a little on a daily basis (Ishak and Al-Deek, 2002). In cloud computing, the need to adequately forecast the traffic stream precisely is necessary for fulfilling clients' needs and benefitting specialist organizations. Various reviews have examined activity stream forecast and 85 % of expectation has been accomplished. However, the work in studying the introduction of the system that can be used to predict traffic using applications frameworks where not successful (Gong et al., 2010; Iqbal et al., 2011; VlahogiANNi et al., 2014; Saripalli et al., 2011).

There are difficulties in obtaining an 100% exact forecast in cloud computing as a continuous procedure. Yingyu and Yingxu, (2008), proposed a period related movement information and a method for doing things that will be utilizing multi specialist frameworks for activity information administration. They utilized CORBA for operator administration. The system activity among the specialists is high in using this approach. Utilizing cloud computing innovation for this work require proper separation and distribution of available capacity. Gong et al., (2010) focus on asset assignment by proposing the calculation of Predictive Elastic asset Scaling (PRESS) model for a couple of information issues. Its primary objective is to measure, in subtle little segments, the changes in application and to portion available assets accordingly. The issue with technique demonstrates that a strategy is not working accurately, strategy created under 5% over-close figure blunder and near zero under-figure mistake, and versatile supply scaling can both diminish to supply (SLO).

Online traffic prediction in the cloud was estimated by using the auto-regression prediction model (Dalmazo et al., 2014). This technique can occasionally create negative outcomes especially when the current drops suddenly (e.g., CPU = 100%, 90%, 10%). This model tries to deliver an estimated Xt of a framework in light of the past qualities Xtaˆ'1, Xtaˆ'2, utilizing the strategy Xt=a1Xtaˆ'1+a2Xtaˆ'2::: The constants (a1, a2 et cetera) are estimated by in the following equation

$$\sum_{i=1}^{Z} aiR(i-j) = -R(J),\ for 1 \leq J \leq Z \qquad (2.\ 1\ )$$

Where: R is the auto-correlation coefficient of the time series and Z is the length of the sample.

Systems that are accessing cloud computing are often very slow in response. Islam et al., (2012) proposed a dynamic window approach using traffic prediction online in the cloud. Their method improved admission and congestion control, the problem with this method is the focus on short dependency on historical data, Figure (2-7) shows a dynamic window approach as a method to predict traffic in cloud computing.

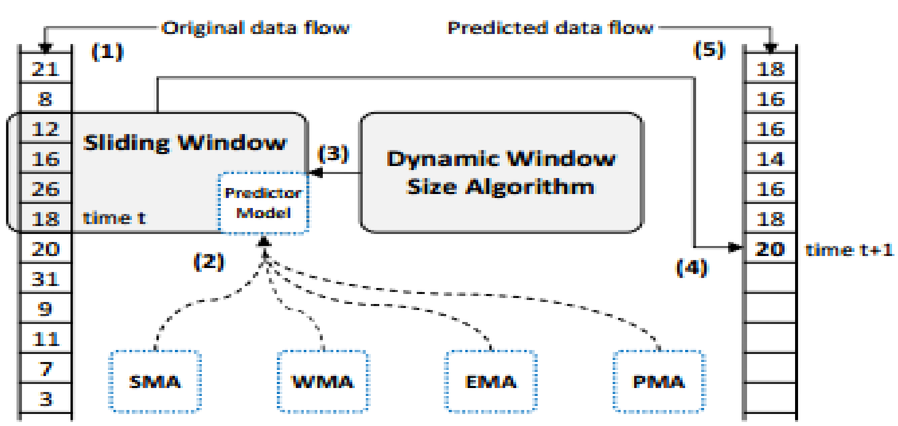

FIGURE 2-7: Window Approach Prediction Method (Dalmazo et al., 2014)

For ongoing factual activity grouping, Nguyen et al.(2012) proposed to arrange web based diversion developments with a minor sliding window in light of a model prepared on various short sub-streams. As indicated by Nguyen et al.,(2012) some tests were exhibited on ID of web based diversion which is an perfectly correct use, but more trials of this test will improve execution.

Saripalli et al. (2011) used forecast load and hot spot detection models for autonomic cloud computing to ascertain the system of movement in cloud computing. They utilized Exponential Moving Average (EMA), and determined that the limits of this techniques is at focuses not extremely distant from the most recent accessible focuses.

The activity expectation device assumes an import part in the conjecture condition. (Zohar et al., 2011) presented an unessential set of PC guidelines. These papers accept that the switches have accumulated information, and they show interest in finding ways to improve the utilization of this accumulated information. Zohar et al., (2011) presented a changing problem adjusting the set of PC instructions for cloud computing in light of a current arrangement of PC guidelines called weighted least connection (WLC) (Ren et al., 2011). The WLC set of PC guidelines appoints undertakings to the hub in view of the quantity of associations that occur for that hub.

This is prepared in view of a correlation of the sum of associations of every hub in the cloud and afterward the movement is separated and offered out to the hub with least number of associations. WLC does not take into account the procedure of painstakingly pondering the capacities of every hub, for example, agreement speed, stockpiling capacity (Ren et al., 2011). The proposed technique is called ESWLC (Exponential Smooth forecast in view of Weighted Least Connection). ESWLC enhances WLC by considering the period arrangement and trials. That is, ESWLC manufactures the choice of doling out a specific occupation to a hub subsequent to requesting a few actions offered out to that hub and getting the opportunity to perceive the hub capacities. ESWLC manufactures the decision in light of the experience of the hub's CPU control, memory, number of systems and the measure of plate space now being utilized. ESWLC then portrays a conceivable hub that is selected in view of smoothing (Ren et al., 2011).

Various machine learning techniques were introduced to look at the nature of the anticipated information and the execution of the strategy. This can be used as a way to apply improved machine learning methods into system activity classification (Suthaharan, 2014). However, the creators may have given little thought to the perplexing issue of space.

Chang et al., (2015), proposed another contributing, mixed approach in applying insight into PC based considerations for use in a specialised framework which is an agreeable complex activity organized for movement forecast flag. As per Chang et al. (2015), the diverse sub issues are separated from the huge scale activity flag control issue, and are taken care of by a brilliant operator with a fluffy nerve-related basic leadership module. The lower-level specialists settle on the choices that are allocated to the more elevated operators. A specialist will take control by accepting a helpful critical thinking approach and obtaining an operator

A multistage, connected process application, which is changing fast processed data to a required input into the multi specialist framework which incorporates reinforcing of learning, weight changes and vigorously refreshing the relations of fluffy utilizing strategy for traffic stream expectations. On the off chance that the cloud is perceived a multi cloud framework, the approach proposed by Min Cheechoy will have the capacity to scale well for a larger networks (Chang et al., 2015).

Tode et al., (2014) expect activity loads take after a cut off Gaussian appropriation and intend to outline a sensible topology to change something. Activity loads characterize a non-straight creation, and a rehashing plan to get the best conceivable connection capacity required for that legitimate topology, subject to the connection flood chance not being past guaranteed.

Mark et al. (2011), trust a nearly indistinguishable concern from Ye et al.)2011), where they expect a multivariate Gaussian activity dispersion together with a relationship (between the requests) grid. Under their multivariate model, the range, containing a specific wanted rate of all conceivable frameworks, is characterized by a general region. The creators build up a Mixed Integer Linear Program (MILP) to make something almost unconceivable so that the cost in pleasing to any movement framework is taken in the oval zone.

The general improvement for managing uncertainty is that of arbitrary programming (Boyd et al., 2011) .This branch of enhancement clarification is much the same as solid and sound streamlining, yet as opposed to having (just) kept down activity, it is expected that the possibility conveyances of the breaking points are known or can be speculated (a number) in light of verifiable information.

Moayedi and Masnadi-Shirazi, (2008) propose a system activity and location show in light of Autoregressive Integrated Moving Average (ARIMA). In this paper, they analyse the traffic flow stream with a specific end goal to isolate from activity distinction rendition. The creators then attempted to autonomously foresee movement. Their work was to make sense of their information and relied on extensive memorable information.

Popa et al., (2013) propose the use of Cicada prediction that ensures bandwidth without requesting the customers to determine their requirments. Obviously the issue with this occasionally anticipates activity incorrectly.

Traffic flow prediction and transfer data gathering devices have been proposed to gather execution data of Hadoop. Nguyen et al., (2013) propose one Hadoop utility to gather cross-layer occasion trails for introduction. To abstain from evolving Hadoop, specialists proposed to disconnect investigation on Hadoop log documents for execution tuning and location (Peng et al., 2014). Be that as it may, Hadoop Watch is centered on anticipating movement requests in view of constant record framework checking.

Wang et al., (2012a) proposed to have an apparatus three-level plan and connected thoughts of specialist based control to arrange movement and transference frameworks. The outline empowers the advancement of a "nearby basic, remote complex" system based plan for minimal effort and high introduction activity frameworks. By tolerating cloud computing innovation-based progressive multi cloud provisioning, the cost can be lessened.

There is a couple of basic ways to deal with anticipating of loads, for example, direct forecast, neural systems and questionable judgment (Duy et al., 2010). The distinction in these forecast strategies is that the models depend on either information focuses or isolated demonstrating. This method for doing things utilize flows to characterize the circumstances between the expectation of the client requirements (between landing times), the duration for a device to facilitate a user request, and the time it takes for the plan to change between its energy conditions (i.e. rest, sit out of gear and dynamic).

In partitioned models, the change state periods are settled. Nonetheless, expecting that the period it takes to change, starting with one state and moving onto the next, is settled may not be valid in all circumstances. For instance, the gear might humiliate over period and accordingly the change may contrast. A green booking set of PC directions incorporated with neural system has additionally been proposed for streamlining assets in the cloud. Neural system forecast approximates the active approaching burden, and fill in as contribution to a green planning technique for turning servers on and off. These judgments work to make it feasible for the quantity of running servers (Duy et al., 2010).

BLINC (Ganapathi et al., 2010) is another approach in light of distinguishing examples of host conduct to isolating and naming activity streams. This is orthogonal of an approach used as a part of Internet activity grouping benchmark apparatus (Netramark) that depends on the behaviour of the movement stream. In future work we would plan to consolidate this strategy to autoregulation coordinated moving normal.

The Internet's availability structure is characterized by Internet Service Provider (ISP) associations by means of the Border Gateway Protocol (BGP), which creates and advances ways for steering messages. LaCurts et al., (2014) concentrate on workload prediction and placement in cloud computing Systems. They proposed Cicada's expectation calculation that makes broken forecasts of system approaches or active movement due to characteristic impediments or inadequate earlier data. The technique will anticipate organized activity and measure and adjust movement forecasts in order to build the execution.

Many sets of computer instructions for the traffic forecasting model have been proposed. The simplest one is old average which has been put into use but lacks quality. ARIMA time series is the general parametric modeling method that was used to predict traffic in many researches and it delivers good accuracy (Prevost et al., 2011).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(A - P)^2 \qquad (2.\ 2)$$

where A is the actual load value, P is the predictive load value and N is the number of samples

Stack Prediction and the Hot Spot discovery technique for autonomic cloud computing for system activity forecast in cloud computing was utilized as well as the Exponential Moving Average (EMA). The limitations of this strategies is anticipated to have a focus not extremely distant from the most recent accessible points (Saripalli et al., 2011).

$$EMA(\vec{S}\,n(ti)) = \alpha * Si + (1-\alpha) * EMA(\vec{S}\,n(ti-1)) \qquad (2.\ 3)$$

where α = 2/n+1 is a constant of smoothing factor; n is the first measure arithmetical mean and the EMA value is initialized of samples.

Deri and Fusco, (2013) utilize hereditary programming configurations to shape a Flexible Neural Tree (FNT) for associated activity expectations. This approach was utilized for a superior comprehension of the principle components of the activity information. In short, what happened to their strategy is that it just records data on minor scale movement forecast limits and can imitate the factual elements of genuine activity measurements. Boto-Giralda et al.,(2010) actualized a parallel extensive impartial system procedure for activity stream expectation in view of a MPI (Message Passing Interface) programming model.

Kalbasi et al., (2012) proposes the Demand Estimation with Confidence (DEC) way to deal with overcoming the precariousness of multi co-linearity in the relapse method. DEC can be iteratively connected to enhance the estimation exactness.

Israil et al., (2015) proposes an Energy Conserving Resource Allocation Scheme with Prediction (ECRASP). In this plan, approaching developments can be anticipated i.e. thick or inadequate. Keeping in mind the end goal is to foresee that creators utilize the idea of exponential smoothing expectation.

As per Tammaro et al., (2012) computing work are considered by their coming and tear despondent circumstances, and additionally a prescient profile of their processing necessities through their movement period. Tammaro et al., (2012) propose that earlier information of the anticipated computing assets required by end-clients is required and inspect a few techniques with various advancement criteria.

Nonetheless, expectation blunders may occur, resulting in one or more processing requests to be dropped. Beloglazov and Buyya, (2012) forms a VM solidification technique that makes use of a speculation related method for doing things that contemplates the impact of co-found VMs to QoS numbers that measure traffic flow prediction. In this procedure, the workload is displayed by assets of a Kalman channel, while the asset utilization profile is evaluated with unseen Markov techniques. The proposed system is tried against SPECWeb 2005.

With regards to today's cloud organizers, this objective puts Cicada a a step in front of research that accepts that clients definitely know these examples in advance (LaCurts, 2014) .

Cicada's prescient approach, in a recorded setting, takes after a few types of ISP movement building. Activity building uses evaluations of movement to indicate flow in a system, and endeavors to anticipate the measure of movement that connections in a system will convey. In any case, Cicada handles a marginally unique issue not covered by other movement designing plans.

Zhang et al., (2003) proposes genuine activity networks from clear estimations, for example, interface stack information. Cicada, then again, knows the correct movement lattices seen in past ages and its issue is to anticipate a future activity framework. Ganapathi et al., (2009) present more propelled machine learning procedures for anticipating asset use. The Convolutional Neural Network (CNN) and the Linear Regression techniques are employed for forecasting. They incorporated a sliding window strategy to mirror the present condition of the plan.

LaCurts, (2014) produced prediction standards based on window sizes and evaluated them with MAPE, PRED (25), RMSE, and R2 prediction accuracy. The CPU utilization data is collected from the Amazon EC2 cloud through the TPC-W benchmark and prediction values are generated with the ECNN and linear regression methods. The prediction values of CPU operation has an 19% error rate without the sliding window and has a minimum of 18.6% error rate when they employ the sliding window.

Mathematical technique also plays an important role in cloud computing prediction which may be used to help determine gain or loss of users or company profits. Blume et al., (2010) propose the movement expectation approach in view of the arithmetical system where perceptions are weighted with Poisson dissemination inside a moving window. They consider the past information by employing a sliding window of size λ and this window is connected by weighting the past perceptions as indicated by the Poisson appropriation with parameter λ. Drop-box follow information is utilized for testing their forecast and Normalized Mean Square Error (NMSE) assessment strategy is used for the error estimation.

Future expectation is one of the critical components in cloud computing since it can help in the arranging process. Smith et al., (2002) predicted a fleeting movement stream by K closest neighbor calculation. We could anticipate the state of the future activity stream through the past changes in activity stream strategy because the activity stream changes is not totally irregular and there are intermittent plans. The little term activity stream can be anticipated by doing a KNN calculation. Islam et al., (2012) proposed an expectation structure that utilizes an arithmetic strategy which can hypothesize the future surge in asset requirements, and, along these lines, empowering proactive scaling to deal with transient blasts workload, in a controlled fashion. For the expectation approach, we alternated to machine learning techniques (e.g. neural network and linear regression) and sliding window systems.

Su and Yu, (2007) proposed a parallel movement stream forecast program for the huge scale road arrange created by Java dialect, Hadoop, GA calculation, and SVM. The proposed calculation, in light of the continuous information, the outcomes expectation measures, and the genuine principles by two parallel techniques, is superior to serial calculation. At the point when the activity stream is fluctuating imperatively, the entire relative mistake of parallel calculation in view of map reduce is moderately steady.

Tomás and Tordsson, (2013) aim at dynamic asset provisioning, misusing flat flexibility. Two versatile half and half controllers, including both responsive and down to earth activities, are utilized to choose the quantity of VMs for a cloud administration to meet the SLAs. The future request is anticipated by a lining system show.

As indicated by Zhang et al., (2010), a key-esteem store is a dormant Service Level Objective (SLO). The proposed middleware finishes high adaptability by utilizing duplication and giving more unsurprising reaction times. An expository model, in view of aligning hypothesis, is displayed to illustrate the connection between the quantity of imitations and the office level, e.g., the portion of solicitations prepared by service level objectives.

Forecasts of system movement is important for some administration applications, for example, asset portion, affirmation control and blockage control. Dalmazo et al., (2014) proposed a dynamic window measure philosophy to join with existing activity expectation systems. This technique is encouraging on the web movement forecast because of its short reliance on recorded information. The greatest hypothetical fluctuation of a given arrangement of information can be assessed from the result of the distinction of its extraordinary qualities, (least esteem to most noteworthy esteem), and the normal, as taken after:

$$\sigma^2 \max = m(Yb - m) - Ya(yb - m) \qquad (2.\ 4)$$

Where $\sigma^2 \max$ acknowledges the minimum, the maximum and the average of the data inside the sliding window, and σ^2 represents the maximum and the average of the data inside the sliding window.

Network traffic prediction has gotten a lot of consideration from established researchers as a way to help administering and overseeing PC systems. Deri and Fusco, (2013) proposed a parallel movement stream forecast strategy for space-time, two-dimensional mix, in view of Support Vector Machine (SVM), however this technique is more reasonable for crisis cases, and not common practice for the ordinary activity condition.

Activities in cloud computing is in disagreement and assets are unmanageable. Shah et al., (2013) proposed a resource allocation algorithm utilizing load balancing of the virtual machine algorithm in cloud computing and the constraint of these calculations is diminishing the reaction time of the virtual machine and not expanding further overheads on the framework and algorithm used to locate the normal response time of each virtual machine; the normal reaction time can come about with the assistance of the equations.

Wang et al., (2012b) describes a procedure that predicts the best steering over a space of movement expectations, comprised of the arched structure of past activity frameworks. This strategy depends on reported time information, which makes it hard to utilize it in a dynamic situation such as cloud computing. Rewards of SMA is, for example, its simplicity, low multifaceted nature and simplicity of utilization.

Liu et al., (2012) consider change discovery strategies for fast system movement. The reason for this work is to create trustworthiness for identifying huge changes in immense information streams with an extensive number of streams. Through a model making use of a Weighted Moving Average (WMA), the calculation appraises the estimation of the following interim, having the capacity to identify Disseminated Dissent of-Administration (DDoS) and sweep assaults. For that, all activity expectations that do not coordinate the reference model are viewed as an oddity.

The exploratory outcome demonstrated that the speed of the strategy was more than two times as fast as the technique, their strategy utilize parallel activity stream strategy in view of SVM (Support Vector Machine), and the test demonstrated that the consequence of the parallel SVM strategy is superior to the BP nonpartisan system method, and when the quantity of parallel hubs is 100, the running time of two thousand connections was 36.48 s (Boto-Giralda et al., 2010) .

The calculation of SOGM (2,1) display for a couple of information issues, its unbiased principle is to give hope for small designing advancements. The new model can accomplish high accuracy contrasting and old-style GM (1,1) and GM (2,1) strategy. Furthermore, it should be versatile to numerous indeterminate framework issues (Di et al., 2012). The model Construct the dim differential condition of SOGM (2,1) model as per the following.

$$ax^{-1}K + X^{0}K + 2^{z}ak = b \qquad (2.5)$$

Where a is the low order parameter and high order parameter respectively and b denotes the control variable.

Here and now activity stream forecast strategy is used in view of the hereditary unbiased systems in cloud computing conditions. Its running proficiency was an additional fourteen times as quick as the consecutive hereditary nonpartisan system method (VlahogiANNi et al., 2014).

Petković et al., (2014) proposed an expectation approach that consolidated Adaptive Neuro based Uncertain Implication Systems (ANFIS) and gathers methodology to foresee the future CPU stack in view of the chronicled information in a network situation. The notable CPU stack information was first partitioned into sub-bunches utilizing the fluffy C-implies grouping. Each sub-bunch was then nourished to neighborhood ANFIS forecast models.

The proper ANFIS bunch is then used to foresee the future esteem of the CPU stack. Expectation with master exhortation and conformal indicators were joined to give execution of prediction traffic to ensure forecasts to network activity demands (Niu et al., 2012) . Their objectives were to develop an expectation strategy that performs only slightly worse than the best arrangement of PC guidelines from a settled arrangement of sets of PC directions and furthermore to give substantially and delivering a great deal with almost no wasted time (or space). The master domain was utilized to finish the primary objective while conformal expectation was utilized for the second objective.

As per (Niu et al., 2012), the clone discovery procedure was a metric based examination method. Asset forecasts, once a clone level was recognized, was done utilizing direct and multi-straight 20-regression. Govindan et al., (2011) investigated and ascertained both static and changing helpful supply provisioning in three diverse activity designs, namely, weekly or standard, large spike and random by utilizing a scoring set of PC guidelines in view of accessibility and cost. The authors established that changing useful supply provisioning outperforms the expensive static provisioning by about 93% in cost reduction. Their valuable prediction method as a way of doing things was based on historical data and the employment of linear moving backward and auto moving backward of the order.

In the accompanying segment we took a look at the technique of traffic prediction in cloud computing. It is exceptionally huge for diagrams and the calculation methods used to foresee movement streams that are like our proposed module and expound their pros and cons, in doing such it will help us demonstrate that there is a requirement for calculation procedure that will anticipate present and future traffic precisely. The accompanying Table 2-1 studies and analyzes points of interest and inconveniences of strategy of system activity expectation in cloud computing.

TABLE 2-1: Pros and cons of prediction algorithms in Cloud computing.

	Pros (advanteges)	**Cons (Disadvantages)**
PRESS (Gong et al., 2010)	• predictions achieve high accuracy	• prediction error frequency results due to space limits
ARIMA (Mark et al., 2011)	• predictions accuracy	• Depends on large historic data
X-Trace (Li et al., 2010)	• perform off-line analysis	• Real-time file system monitoring.
ESWLC (Ren et al., 2011).	• More accurate results than WLC	• Complicated, Prediction algorithm requires existing data and has long processing time

	Pros (advanteges)	Cons (Disadvantages)
DWP (Islam et al., 2012)	• congestion control	• Focus on short dependency on historical data
(LaCurts et al., 2014)	• Guarantees Bandwidth	• Faulty prediction.
ANN (Marston et al., 2011)	• congestion control	• Faulty prediction.
EMA (Kalbasi et al., 2012)	• match representative SaaS	• Faulty prediction.
FNT (Deri and Fusco, 2013)	• predictions accuracy	• small-time scale

Traffic prediction in cloud computing plays an important role in IT Service providers and users that use the services, in order to predict traffic accurately for current and future to help providers and users to meet the service level agreement (SLA) and Operation Level Agreement (OLA). In the following Table 2-2 we study and compare the technique of network traffic prediction in cloud computing.

TABLE 2-2 Traffic Prediction Technique in Cloud computing

Author	Method	Merits	Demerits
Wang et al., (2014)	Energy Conserving Resource Allocation Scheme with Prediction (ECRASP).	Accurate prediction	Error Associated with Estimation of Job Execution Times.
Roy et al., (2011)	Just-in-time Resource allocation	Cost effective	1. Prediction error 2. Use of recursive data 3. Structures.
Kalbasi et al., (2012)	Demand Estimation with Confidence (DEC)	Improve the estimation accuracy.	Network Delay.
Tammaro et al., (2011)	Investigate several algorithms	Cost effective, Less delay	Prediction errors
LaCurts et al., (2014)	Cicada's workload prediction algorithm	Cheaper to execute, Improve the prediction rating accuracy,	It is computationally intensive, highly volatile in nature

Author	Method	Merits	Demerits
		Predict the future accurately	
Gong et al., (2010)	Predictive Elastic Resource Scaling algorithm	Simple to use and cheaper to execute, Minimize resource waste.	The prediction error, Faulty prediction
Yoon et al., (2015)	Advanced machine learning techniques	Prediction accuracy	Long forwarding chains, Delayed user experiences.
Lopes Dalmazo et al., (2013)	Statistical model	Response time. Improved resource utilization, prediction accuracy.	Error Associated with Estimation of Job Execution Times.
Xu et al., (2003)	K nearest neighbor algorithm	Cost effective	Not suitable for interactive real time applications
Islam et al., (2012)	Statistical model	Shorter Prediction time	Handle temporal bursty workload

Author	Method	Merits	Demerits
Li et al., (2010)	Parallel traffic flow prediction program	Prediction accuracy	Not cost effective
Kouki et al., (2015)	Queuing-network model.	Simple to use and cheaper to execute	Prediction error

In the previous section we explored some traffic prediction techniques that are used to predict traffic in the current situation. It could be seen that none of the existing traffic prediction techniques are capable of predicting traffic 100 % accurately. There are Several methods that service providers will consider before implementing the solution to the current traffic Flow, Table 2-2 studies and compares the techniques used for network traffic therefore there is a need for an algorithm for the prediction of accurate traffic in cloud computing .

Qibo et al., (2009) propose a system activity expectation and irregularity discovery strategy in view of Auto-Relapse Integrated Moving Average (ARIMA). They contend that they disintegrate the information stream, keeping in mind the end goal to detach peculiarities from typical movement variety. The authors then tried to predict anomalies independently from normal traffic. The following section focuses on purposes, limitations of the module, and the data set of some of the algorithms used to predict traffic in cloud computing. In doing so we can easily identity the advantages and disadvantages of some technique use to predict traffic in cloud computing. In Table 2-3 limitations of the techniques that we saw in the literature review and the data set that was used for each and every method can be seen.

TABLE 2-3: The Main Purposes of Traffic prediction Technique in Cloud computing

r

Authors	Techniques	Data set	Purpose	Limitations
Boto-Giralda et al., (2010)	Neural network	Data set from B-ISDN.	Decrease noise video traffic in term of network traffic prediction	Time delay
VlahogiANNi et al., (2014).	GARCH	Abilene backbone internet	To predict and capture the bursty nature of Internet traffic	Prediction error
Dinh et al., (2013)	D-BLRNN	Bell core data set	To improve the BRNN approach for network traffic prediction	Prediction error
Ganapathi et al., (2009)	Elman Neural Network	Simulation	Long –rang internet traffic prediction	The prediction error, Faulty prediction

Shah et al., (2013)	MLP	From backbone internet	To predict the internet traffic over IP network	Faulty prediction
Govindan et al., (2011)	ARIMA and GARCH	University of Calgary	To create conditional mean and conditional variance model for prediction internet traffic	Time delay
Petković et al., (2014)	ARIMA/GARCH	From University of Auckland’s Internet uplink in different	To predict busty nature of network traffic	Time delay
Liu et al., (2012)	Neural Network Ensemble	From internet service provider on based TCP/IP protocol	To predict error from internet traffic	Error Associated with Estimation of Job Execution Times.
Petković et al., (2014)	ANFIS	From backbone internet over TCP/IP	For modeling and predicting internet traffic	Network Delay.

Zhang et al., (2010)	BPWNN	Collected network dataset from real link	Improving drawback of BPNN technique in network traffic prediction	Handle temporal bursty workload
Qibo et al., (2009)	ARIMA and MLANN	Web server of Guangzhou network center	To predict error from network traffic	Delayed user experiences
Mark et al., (2011)	Neural network and ARIMA	Free stats and Fotoblog websites	To predict short-range and long-rang self-similar network traffic	Time delay
Niu et al., (2012)	Decomposed model	Backbone router of NSFNET	To predict long-rang network traffic	Faulty prediction
Beloglazov and Buyya, (2012	FARIMA and neural network	CRAWDAD website	To forecast wireless network traffic for improving quality of service (QoS).	Delayed user experiences

Xu et al., (2003)	k-Factor ARMA(autoregressive moving average)	Collected dataset from different real network data like MPEG, VIDEO, JPEG, INTERNT, and ETHERNT.	Predicting multi scale high speed network traffic by using different dataset.	Not suitable for interactive real time applications
Smith et al., (2002)	APM	Certain mobile network of Heilongjian province in chain	To predict mobile network traffic	Time delay
Wang et al., (2012b)	SARIMA	Public safety trunked radio network	To predict call from radio network	Faulty prediction
Deri and Fusco, (2013)	AARIMA	Start ware MPEG and Bell core	For predicting long-rang-dependant internet traffic	Minimum execution time.
Deri and Fusco, (2013)	EPTS	ECNU data set	Predict network traffic for managing bandwidth	Faulty prediction

Smith et al., (2002)	K-means	NS2 simulation over TCP and UDP protocol	Avoid congestion of mesh wireless network in long- rang dependence	Faulty prediction

2.22. Algorithm Evaluation

In cloud computing, movement stream forecast assume a noteworthy part in today's shrewd registering of frameworks. As should be obvious in Table 2-3 that the calculation applied have demonstrated the constraints, yet some techniques are having comparative restriction on more than four appearances, for instance Auto-Regressive Integrated Moving Average (ARIMA) indicates to having impediment of time deferral and minimum execution time where Artificial Neural Network systems (ANNs) have a confinement of delayed client encounters in such a manner that we contrast the two techniques agreeing with our finding in the literature. In doing so, we will actualize the two strategies and think about the outcomes and see which strategy can anticipate the movement extremely well.

2.23. The Importance of traffic prediction in cloud computing

In cloud computing clients care only about how the system functions and responds. They are not interested in the background process of the applications, as long as the user can perform their function without wasting time or waiting for the system to respond. The work of the service provider is to keep the service level agreement. The most important reason for why traffic prediction is very important are.

- Better traffic flow in cloud computing;

- Congestion Control;
- Minimise cost usage and increase productivity;
- Increase system respond time;
- Better planning information;
- Greater user acceptance;
- Avoidance of system bottleneck; and
- Increased business activity.

2.24. Architecture of a traffic prediction system

In this research, the following problems are addressed in cloud computing traffic flow:

- The problem of incorrect traffic prediction in cloud computing;
- Traffic flow congestion management;
- System delay due congestion control;
- Traffic flow workload management;
- Response time; and
- Reduce prediction error.

2.25. Summary of chapter 2

In this section, we considered and presented cloud computing, we detailed the critical layers and advantage of cloud computing which is Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). Apart from that we took a look at three types of cloud computing (private, public and hybrid cloud) and their points of interest as well as their weaknesses. We showed that there is current and future challenges in cloud computing. In this chapter we also summarized traffic prediction techniques and their limitation in cloud computing and we compared algorithms that have small limitations. We concluded this chapter with the architecture of traffic prediction systems. The following chapter focuses on methodology.

CHAPTER 3

3. Methodology

The traffic flow prediction in cloud computing proposed in the literature shows that Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) could be appropriate models for traffic prediction in cloud computing. This study will produce two traffic flow prediction techniques that can be evaluated and recommended to be used by any service provider. There is a need to predict traffic flow accurately in order to assign the resources according to the need of the customer or client. This section provides detail information about algorithms used to predict traffic in cloud computing. This method that can be adapted in any amount of traffic according to the need of the clients. Cloud computing has created huge opportunities in technology companies such as Amazon and Microsoft, who have huge server farms with dedicated machines accessible for clients to lease. Their clients vary from individuals to large organizations with workloads running from short, profitable supply-concentrated employments to long-running client administrations.

The cloud is considered to be a more reliable and consistent system than in-house IT infrastructure. In this case, if cloud computing is to fulfill its drive, there should be a clear understanding of the diverse issues involved in the innovation, both from the perspectives of the suppliers and the general population (who utilize an item or administration) (Marston et al., 2011). This study proposes a suitable calculation to foresee movement in cloud computing and suggests how an ARIMA and/or ANN can be utilized to anticipate activity in cloud computing and the apportion the assets.

This simulation proposes Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) to predict traffic in Cloud computing.

3.1. Auto-Regressive Integrated Moving Average (ARIMA)

Auto-Regressive Integrated Moving Average (ARIMA) is a model using smallest squares estimation of AR coefficients which can be precisely computed from autocorrelations in a solitary cycle.

According to Khashei et al.(2009), ARIMA is divided into three stages, estimate, forecast and ?? which are summarized below:

- In the ID stage you utilize and recognize explanations to determine the reaction arrangement and distinguish the suitable ARIMA models for it. The distinguished proclamation examines time arrangement which is to be utilized as a part of later articulations, conceivably differentiating them, and determines autocorrelations, opposite autocorrelations, halfway autocorrelations, and cross connections. Unmoving tests can be finished to make sense if differencing is essential. The examination of the recognized articulation more often than not recommends at least one ARIMA model that could be a fit. Choices enable you to test for unmoving and uncertain/not unmistakable ARIMA arranged ID.

- ARIMA problem recognizing stage: you use the gauge proclamation to determine the correct ARIMA model to fit the predefined description, and to configure the rules for that model. The gauge proclamation additionally creates distinctive measurements to help you judge the effectiveness of the model. Significance tests for rules determine whether some of the terms in the model might be pointless. Integrity of fit insights help in contrasting this model with others. Tests for repetitiveness indicate whether the additional measurements contains more data that may be suitable for a more comprehensive model. On the off chance that the diagnostic tests (to get data) indicates issues with the model, attempt another model, after which the significance and integrity tests should be repeated.

- The estimation stage is used to determine future estimations of the time arrangement and to determine certain timeframes for these measures from the ARIMA model established by the past/preceding appraisal articulation.

These three stages are applied in the following ARIMA models and the ARIMA formula is written as:

$$W_t = \mu + \frac{\theta(B)}{\phi(B)} a_t \quad (3.1)$$

Where:

t Indexes time

w_t Response series Y_t or a difference of the response series

μ The mean term

B The backshift operator; that is, $BX_t = X_{t-1}$

$\phi(B)$ The autoregressive operator, represented as a polynomial in the back shift operator: $\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p$

$\theta(B)$ The moving-average operator, represented as a polynomial in the back shift operator: $\theta(B) = 1 - \theta_1 B - ..... - \theta_p B^q$

a_t The independent disturbance, also called the random error.

The arrangement W_t is figured by the distinguish explanation and is the arrangement handled by the gauge articulation. Thus, W_t is either the response series Y_t or a difference of Y_t specified by the differencing operators in the identify statement.

For simple (non-seasonal) differencing $W_t = (1-B)^d Y_t$. For seasonal differencing $W_t = (1-B)^d (1-B^s)^D Y_t$, where d is the degree of non-seasonal differencing, D is the degree of seasonal differencing, and s is the length of the seasonal cycle.

Model Constant Term:

The ARIMA model can also be written as:

$$\phi(B)(W_t - \mu) = \theta(B)a_t \tag{3.2}$$

Or

$$\phi(B)W_t = const + \theta(B)a_t \tag{3.2}$$

Where:

$$const = \phi(B)\mu = \mu - \phi_1\mu - \phi_{2\mu} - ..... - \phi_p\mu$$

Thus, once an autoregressive operator and a mean term are both included in the model, the constant term for the method can be represented as (B). This value is issued with the label "Constant Estimate" in the estimate statement output.

Therefore: Notation for Transfer Function Models

$$W_t = \mu + \sum_i \frac{_{wi}(B)}{\delta_i(B)} B^{ki} X_{i,t} + \frac{\theta(B)}{\phi(B)} a_t \tag{3.4}$$

Where*:*

$X_{i,t}$ The ith input time series or a difference of the ith input series at time t

$_{ki}$ The pure time delay for the effect of the *ith* input series

$_{wi}(B)$ The numerator polynomial of the transfer function for the *ith* input series

$\delta_i(B)$ The denominator polynomial of the transfer function for the *ith* input series.

The model can also be written more compactly as

$$W_t = \mu + \sum_i \Psi_i(B) X_{i,t} + n_t \tag{3.5}$$

Where:

$\psi_i(B)$ The transfer function weights for the *ith* input series modeled as a ratio of the! And polynomials: $\psi_i(B) = (_{wi}(B)/\delta_i(B))B^{k_i}$

n_t The noise series: $n_t = (\theta(B)/\phi(B))a_t$

ARIMA models are sometimes expressed in a factored form. This means that the $\phi, \theta, _{w,}\ or\ \delta$ polynomials are expressed as products of simpler polynomials. For example, we could express the pure ARIMA model as:

$$W_t = \mu + \frac{\theta_1(B)\theta_2(B)}{\phi_1(B)\phi_2(B)} a_t \tag{3.6}$$

Where:

$\phi_1(B)\phi_2(B) = \phi(B)$ is the autoregressive operator, represented as a polynomial in the back shift operator: $\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p$

$\theta_1(B)\theta_2(B)=\theta(B)$ is the moving-average operator, represented as a polynomial in the back shift operator: $\theta(B)=1-\theta_1 B-.....-\theta_p B^q$

3.2. Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are a group of models given incredible considerations such as nerve-related/cerebrum related systems and are utilized to determine a number or mechanics that can rely on an extensive number of data sources and are generally vague (Samanta and Al-Balushi, 2003). Artificial Neural Network (ANN) is broke down in The back engendering calculation utilized as a part of layered nourish forward ANNs.

This implies that the nerve cells are sorted out in layers, and send their signs "forward", and after that the errors are spread in reverse. The possibility of the back spread arrangement of PC directions is to decrease this error, until the ANN takes in the preparation information. The preparation starts with irregular weights, and the objective is to improve or change to fit new conditions so that the blunder will be practically nothing.

Artificial neural networks (ANNs) function can be written as:

$$y=\theta\left(\sum_{j=1}^{n} w_j x_j - u\right) \tag{3.7}$$

Where*:*

$\theta(.)$A Unit step Function at 0

W_j The synapse weight associated with the jth input

X_j Weighted sum of its n input signals, e.g. $x_j, j=1,2,.....,n,$

u Simplicity of notation or another weight,

$w_0 = -u$ Attached to the neuron with a constant input $x_0 = 1$

3.3. Data Collection

Traffic flow data in cloud computing was gathered from 01 April 2011 to 30 April 2017 (Columbus, 2017). The data is introduced in Exabyte (EB). The drive for the research is realistic in following the most recent monetary, mechanical, and traffic developments to take care of social developments and their affect on how organizations use programming to interface with and serve their clients. What makes every one of these improvements so alluring is the way trust is getting calculated into these level supporting surfaces and the part it is playing in informal communities.

3.4. Evaluation Measures

Assessment measures are the numerical conditions which are used to characterize errors amongst genuine and anticipated qualities. The contrast amongst real and anticipated qualities demonstrates how well the model has acted. The fundamental thought behind the forecast technique is for foreseeing the result and this ought to influence the execution and of the model (Almeida et al., 2015). These error measures are vital in the expectation of future activity, as the estimations taken by these measurements, exceedingly impact the future arrangement of the associations. As indicated by Almeida et al., (2015) any assessment that measures the errors ought to have five fundamental qualities which are: legitimacy, simple to translate, unwavering quality, satisfactory and must have a numerical condition.

The following quality measures are explained below:

- Validity states to what extend the measures are well-founded. Legitimacy alludes to whether the error measures really measures what it plans to measure. The metric ought to gauge the outcomes that can identify with the offered information. For instance, if the metric is indented to measure a double frame, then matching outcomes will be considered as legitimate yield, generally legitimacy of the metric can be interrogated. Legitimacy likewise alludes to the legitimacy of the estimation, i.e. how true the error measured by a mistake metric is.

- Easy to Understand alludes to disentanglement of the metrics. The metrics ought to be straightforward and stay away from complex problems. The specialists abstain from utilizing complex numbers to measure things, particularly if expectations are complicated and metrics have the capacity to clarify and persuade choices, which is less demanding if the metrics is straightforward.

- Reliability is the equity and uniformity of the mistake metric to quantify error utilizing a similar estimation strategy on the same subject. On the off chance that rehashed estimations are taken and each time the outcomes are exceedingly steady or even indistinguishable, then there is a high level of forecast, yet in the event that the estimations have extensive contrasts, unwavering quality is low. The error metrics is dependable when it will clearly gauge a similar outcome each time it is given similar information.

- Respectability alludes to the capacity of error metrics and its estimation to be spoken to in a straightforward framework.

- Arithmetical conditions suggests that the error metrics can be spoken to as numerical conditions. Numerical conditions are mostly normal and straightforward as it clarifies the type of error numbers that measure things.

3.5. Selected Error Metrics

In activity data, to deliver precise traffic expectations, we train the information on both strategies, on the premise of dependability and wide utilization, after execution (mistake) measuring are chosen to quantify how exact the traffic forecast is.

- Mean Absolute Percentage Error (MAPE)

3.6. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) is one of the most common and popular error metrics for prediction. MAPE calculates the mean of absolute percentage error which is easy to recognize and calculate. The MAPE is represented by the following equation:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{D_t - F_t}{D_t} \tag{3.8}$$

Where D_t is the actual value and F_t is the predicted value

3.7. System Implementation Plan

The following algorithms are going to be experimented:

- Design the procedure of Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs).
- Traffic forecast utilizing Auto-Regressive Integrated Moving Average (ARIMA).
- Traffic forecast utilizing Artificial Neural Networks (ANNs).
- Evaluation measures between Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs).
- Measure the accurate movement expectation and contrast and different researchers.

3.8. Summary of chapter 3

This chapter proposed Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs) to anticipate precise activity expectation and we outlined the data collection taken by evaluation measures, which is Mean Absolute

Percentage Error (MAPE), to gauge the difference between real and anticipated values to help show how well the model has performed. We closed the section by briefly discussing the system implementation plan for this research. The following chapter concentrates on design, experimentation, results and discussion.

CHAPTER 4

4. Design, Experimentation, Results and Discussion

This chapter presents strategy as well as the results of the different experimental tests run for determining prediction competency of the two algorithms techniques, namely Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs). The objective of this section is to present and evaluate the accuracy of the selected techniques in current and future traffic prediction. In addition, the presence of data observed (Columbus, 2017), is considered when discussing prediction. The results include both training and test datasets that was used to build up the traffic prediction system. We are going to use Matlab tools and the parameters used to obtain these results have already been presented in chapter 3. Similarly, the evaluation metrics has also been defined in the previous section (3.4).

The crucial purpose of these experiments are:

- Predict traffic using Auto-regressive integrated moving average (ARIMA);
- Predict traffic using Artificial Neural Networks (ANNs)**;**
- Evaluation Measures; and
- Measure accurate traffic prediction and compare with other scholars' findings.

4.1. Design of Auto-Regressive Integrated Moving average (ARIMA)

For the traffic prediction to be predicted accurately, the ARIMA technique will be used to reduce prediction errors in a step by step fashion. The following are the steps ARIMA takes to produce the output of the results. Firstly the proposed module will perform step 1 which is identification of the model, followed by step 2, the estimation of the parameters and moving on to step 3 which is validation of the model. From step three (3) the model will determine if the module is acceptable and will then proceed to step 4 for prediction. Figure 4-1 shows an ARIMA prototype for traffic prediction in cloud computing.

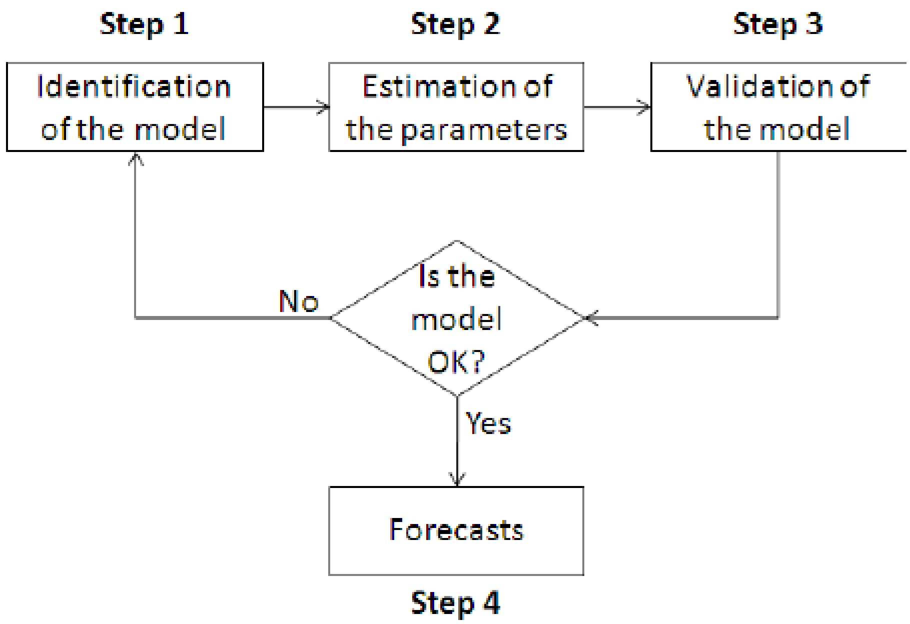

FIGURE 4-1: ARIMA design for traffic prediction (Khashei et al., 2009)

4.2. Design of Artificial Neural Networks (ANN)

The Artificial Neural Networks (ANN) technique will be used to step by step predict traffic in cloud computing. The following are the steps that ANN will run until they produce the prediction results. Firstly ANN will accept the input values, after the acceptance of the input values it will issue synaptic weights followed by calculations. ANN will continue to calculate the activation functions and produce the prediction output. Figure 4-2 illustrates an ANN prototype for traffic prediction in cloud computing.

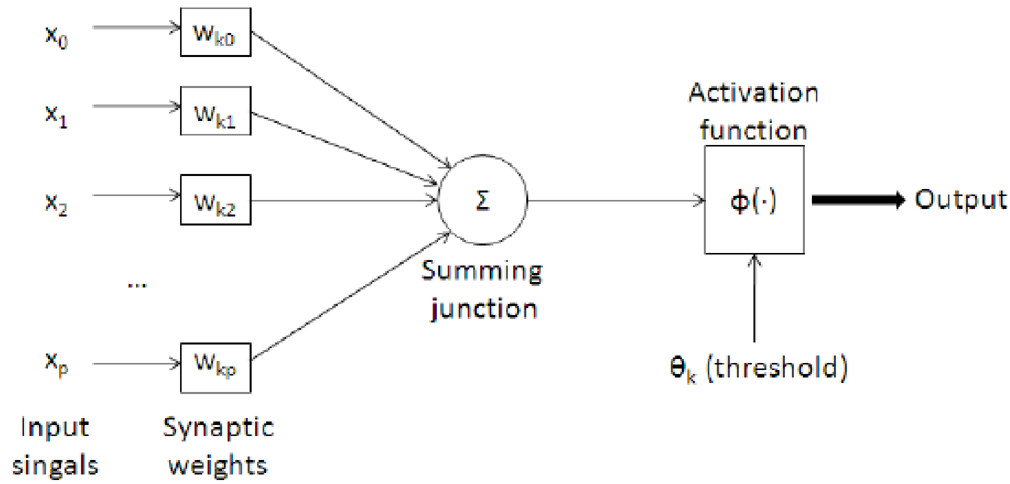

FIGURE 4-2: ANN design for traffic prediction (Samanta and Al-Balushi, 2003)

4.3. Simulation Experiment and Results

The first step of the experiment is to use actual data observed (Columbus, 2017) The dataset is obtained from the Center for Applied Internet Data Analysis (CAIDA) and is used to train, validate and predict the traffic flow. For both algorithms traffic flow data and the period that the data was observed are presented for both training and test datasets.

4.4. Auto-Regressive Integrated Moving Average (ARIMA)

This algorithm is used to analyse the current traffic and predict the future traffic in cloud computing, Figure 4-3 presents a graphical picture of the current and predicted output values where ARIMA shows what will take place in the future and put value measures separately. The training and test space shows accurate prediction values as you can see in Figure 4-3. All training value results are similar to the test results. In fact, the results shows 98 % accuracy. Figure 4-3 demonstrates the traffic flow prediction (measured in Exabyte) per year.

FIGURE 4-3 Comparison of actual data vs predicted data on traffic flow using ARIMA from 2012 to 2019

The Auto-Regressive Integrated Moving Average (ARIMA) technique shown in Figure 4-3 proves that this method can predict traffic very well and can be adopted by service providers and clients to implement and process the cloud computing technology and enjoy the benefits of cloud computing . In order to maximise the benefit of cloud you need to know and plan your data and time usage of the system. The following section elaborate on the traffic flow data, ARIMA traffic flow and prediction from 2012 to 2019.

4.5. Actual Data Flow (2012 to 2017) vs Predicted Data flow (2013 to 2019) using ARIMA

Prediction of traffic from 2012 to 2019 indicates in Figure 4-3 how accurate the ARIMA model is and can be recommended to predict the future. According to our experiment the model shows accurate results therefore service providers will be able to plan and manage their clients according to their requirements. Figure 4-4 shows how close the prediction and actual data measured.

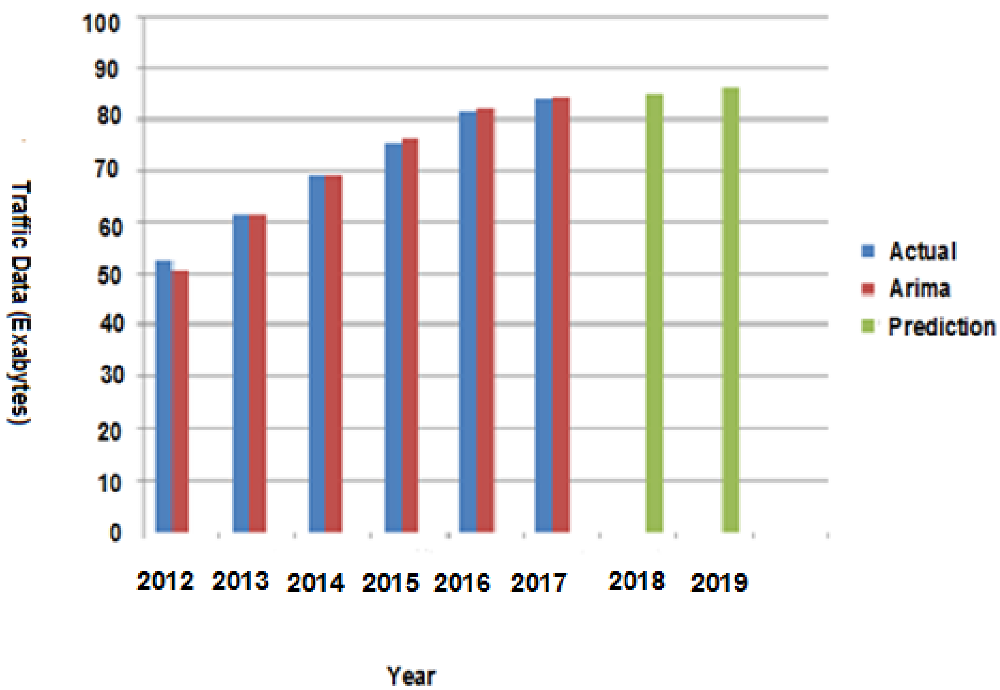

Figure 4-4: Actual Data vs ARIMA vs Predicted Data

4.6. Evaluation Measure of ARIMA (2011 to 2018)

To prove how accurate Auto-Regressive Integrated Moving Average (ARIMA) is, we are going to use Mean Absolute Percentage Error (MAPE) to evaluate the errors between actual and predicted values. The following is a calculation of first year (2012).

Prediction errors for ARIMA from 2012 to 2018 are demonstrated below. We may draw conclusion that ARIMA in the year 2012 was not performing very well but it recovered in the year 2013. The method improved by almost 80 percent and in the year 2014 we saw an increase in error prediction due to congestion control. Lastly we saw a very close prediction in the year 2017. We will use the following MAPE formula to calculate how accurate the predicted values are compared to the actual value.

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \frac{D_t - F_t}{D_t} \qquad (4.\ 1)$$

Where D_t is the actual value and F_t is the predicted value

Table 4-1 demonstrate the summary of errors between actual and predicted data from the year 2012 to 2018 using MAPE. Data is measure in Exabyte per year.

TABLE 4-1: Traffic flow prediction error for ARIMA from 2012 to 2019

Exabyte(ZB) Per Year	Actual data(Dt)	Predictions(Ft)	Errors (Dt-Ft)
2012	52.4	51.03	1.37
2013	61.6	61.4	0.2
2014	69.2	69.01	0.19
2015	75.3	76.3	-1
2016	81.5	82.3	-0.8
2017	84.3	84.4	-0.1
2018	n/a	85.1	n/a
2019	n/a	86.4	n/a
MAPE			-0.023333333

Table 4-1, demonstrates traffic flow prediction error for ARIMA from 2012 to 2018 and also shows how accurate ARIMA technique was by demonstrating MAPE value of -0.23 to achieve our main objective of measuring traffic flow prediction in cloud computing for the different times of the year. We will look at the single year, 2013, to see how data was adopted and predicted during the year. In doing so we will be able to observe if Auto-Regressive Integrated Moving Average (ARIMA) can be used to predict traffic on a daily basis or is it performing well if we are predicting on a yearly basis. The following section focuses on monthly basis for the year of 2013, we observe that actual data and predicted data can be determined for every month.

TABLE 4-2: Traffic flow data and ARIMA prediction data from Jan 2013 to Dec 2013

Observations	Actual Data(Dt)	Prediction (Ft)	Errors(Dt-Ft)
Jan-13	4.3	3.8	0.5
Feb-13	6.1	5.2	0.9
Mar-13	5.1	4.9	0.2
Apr-13	6.3	4.1	2.2
May-13	4.8	5.8	-1
Jun-13	5.5	6.9	-1.4
Jul-13	7.9	5.3	2.6
Aug-13	4.9	4.2	0.7
Sep-13	4.4	5.2	-0.8
Oct-13	5.1	4.9	0.2
Nov-13	6.1	5.9	0.2
Dec-13	7.3	5.2	2.1
MAPE			0.533333333

The traffic flow on a monthly basis is measured in Exabyte's and observed traffic flow in the year 2013 was inconstant as you can see from Table 4-2 where the minimum traffic data was 3.8 Exabyte in January 2013 and the maximum data was 6.9 Exabyte in July 2013, at the same time the predicted data for ARIMA in the year 2013 was 1.8 Exabyte in January and the maximum prediction for ARIMA was 5.1 in November 2013. Figure 4-5 compares actual data and ARIMA data for the year 2013.

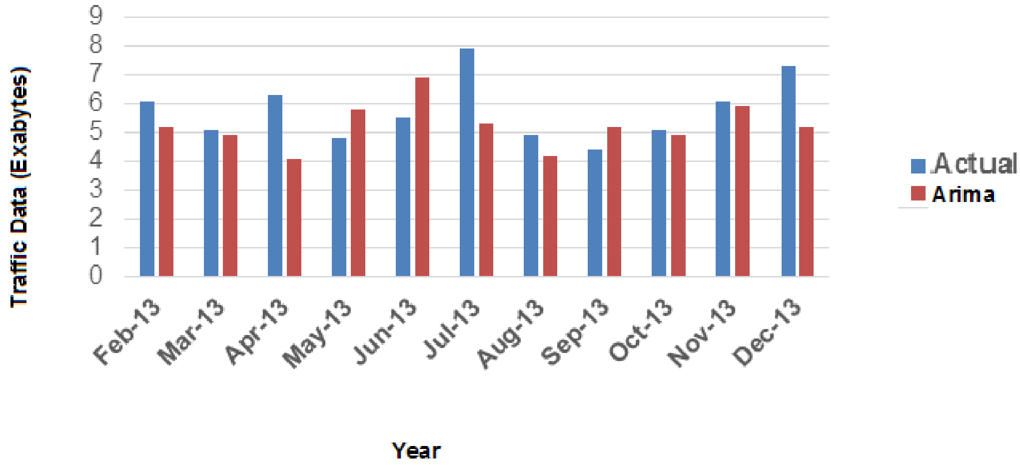

FIGURE 4-5: Actual data vs predicted data flow using ARIMA for 2013

4.7. Artificial Neural Networks (ANNs)

The second step for our experiment is to use the same actual data observed and the following algorithm is used to analyse the current traffic and predict the future traffic in cloud computing (Columbus, 2017). Figure 4-6 is a graphical representation of the actual and predicted throughput values for traffic flow data, ANN prediction and evaluation measures respectively. The training and test interval are 89 % accurate as you can see in Figure 4-6. All training value results are similar to the test results. In fact, the results are 89 % accurate and the data is measured in Exabyte (EB) per year where the MAPE shows the different errors among the actual values and the predicted values.

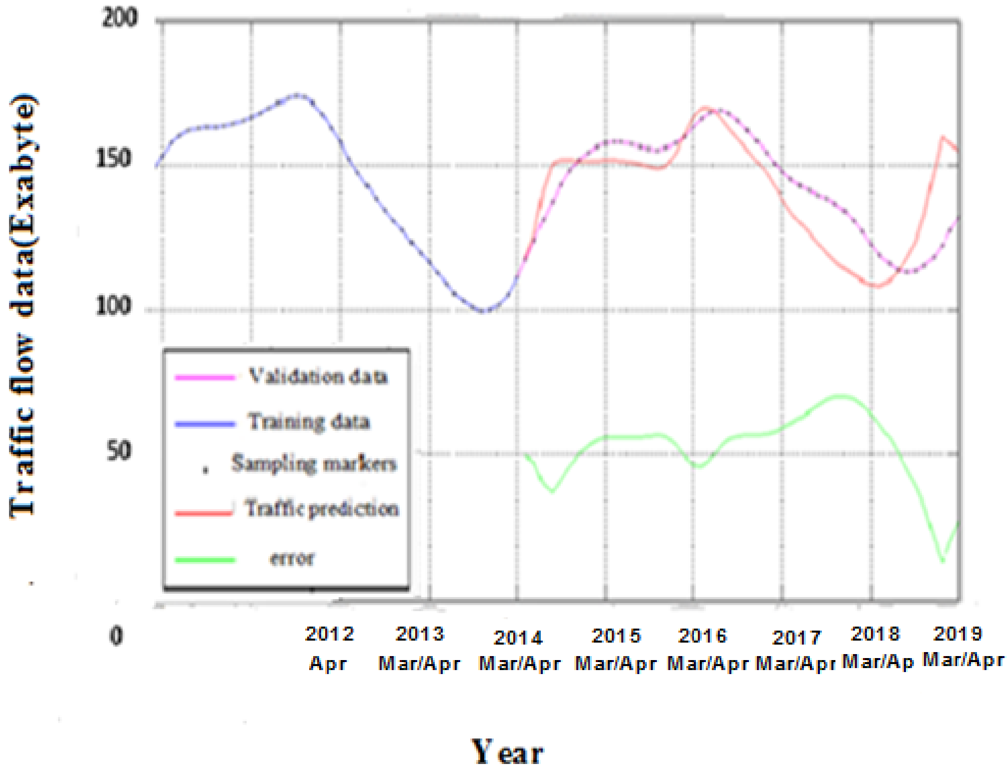

FIGURE 4-6: Traffic prediction using ANN

The Artificial Neural Networks (ANNs) technique shown in Figure 4-6 proves that this method can also be used to predict traffic, the next step is to measure how accurate ANN method is. The following section demonstrates errors shown during the implementation of the Artificial Neural Networks (ANNs) technique.

4.8. Evaluation Measure of ANN (2011 to 2018)

To prove how accurate the Artificial Neural Networks (ANNs) technique is, we are going to use Mean Absolute Percentage Error (MAPE) to evaluate the errors between actual and predicted values and the following is a calculation of errors from 2012 to 2018.

Table 4.3 shows the overall traffic flow prediction error for ANN to determine how close the prediction was to the actual data. MAPE was used to determine the difference between the actual data and the predicted data. In the year 2018 the technique shows accurate prediction with 0.2 percent accuracy compared to the other years. Even if the traffic volume was very high, this model produced accurate predictions. In our view the method can perform better in the overall prediction. The data is measured in Exabyte (EB) per year.

TABLE 4.3: Traffic flow prediction error for ANN

Exabyte(ZB) Per Year	Actual data(Dt)	Predictions(Ft)	Errors (Dt-Ft)
2012	52.4	50.2	2.2
2013	61.6	60.2	1.4
2014	69.2	68.9	0.3
2015	75.3	74.9	0.4
2016	81.5	81.2	0.3
2017	84.3	84.1	0.2
2018	n/a	85.3	n/a
2019	n/a	86.4	n/a
MAPE			-0.023333333

To achieve the main objective of measuring traffic flow prediction in cloud computing on a monthly basis we will take the year 2013 and predict the traffic by using Artificial Neural Networks (ANNs). The data is measured in Exabyte's per month. Table 4-4 demonstrate the traffic prediction per month.

TABLE 4-4: Traffic flow data and ANN prediction from Jan 2013 to Dec 2013

Observations	Actual Data(Dt)	Prediction (Ft)	Errors(Dt-Ft)
Jan-13	4.3	3.8	0.5
Feb-13	6.1	2.2	3.9
Mar-13	5.1	3.9	1.2
Apr-13	6.3	5.1	1.2
May-13	4.8	4.8	0
Jun-13	5.5	5.9	-0.4
Jul-13	7.9	4.3	3.6
Aug-13	4.9	3.2	1.7
Sep-13	4.4	6.2	-1.8
Oct-13	5.1	5.9	-0.8
Nov-13	6.1	2.9	3.2
Dec-13	7.3	4.2	3.1
MAPE			1.283333333

Traffic flow in the year 2013 shows inconstancy from January to December where we see lower traffic flow in January and high volume of traffic in July. However traffic data may fluctuate according to the flow of traffic. Overall the prediction shows promising results in the year 2013, as shown in May that traffic prediction was accurate (4.8 Exabyte) followed by February where we see the ANN technique producing poor results. However, in July where there was a larger volume of traffic flow and we saw the ANN method producing poor results due to high volume. This is the first limitation of this method. Figure 4-7 compares actual data and ANN data for the year 2013.

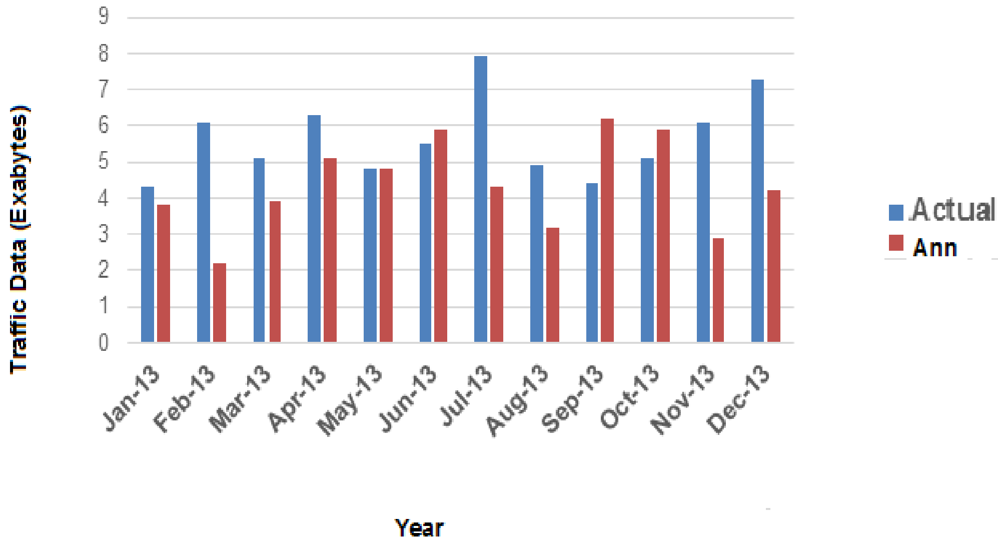

FIGURE 4-7: Actual data vs predicted data flow using ANN for 2013

4.9. Comparison of the two techniques (ARIMA and ANN)

The following section presents the overall results of several experimental simulations to determine which technique is more accurate to predict traffic between Auto-Regressive Integrated Moving Average (ARIMA) and the Artificial Neural Networks (ANNs) model. Figure 4-7 shows actual data and traffic prediction between two models as well as future prediction for the coming two years which is 2018 and 2019. From Figure 4-8 it can be observed that the two models can predict traffic very well. As is shows in Figure 4-8 ARIMA outperforms ANN overall, however ARIMA can produce poor prediction when predicting high volume of data. At the same time, ANN can manage both low and high volumes of Data. The study of their effects on the traffic prediction could be useful for the traffic measurements, traffic prediction as well as network management and control. The ARIMA and ANN traffic prediction models can be used to build effective crowding and blockage control plans, e.g. changing frequency, setting apart, distributing and control.

Figure 4-8 also shows that these sets of computer instructions can be used for future predictions and planning.

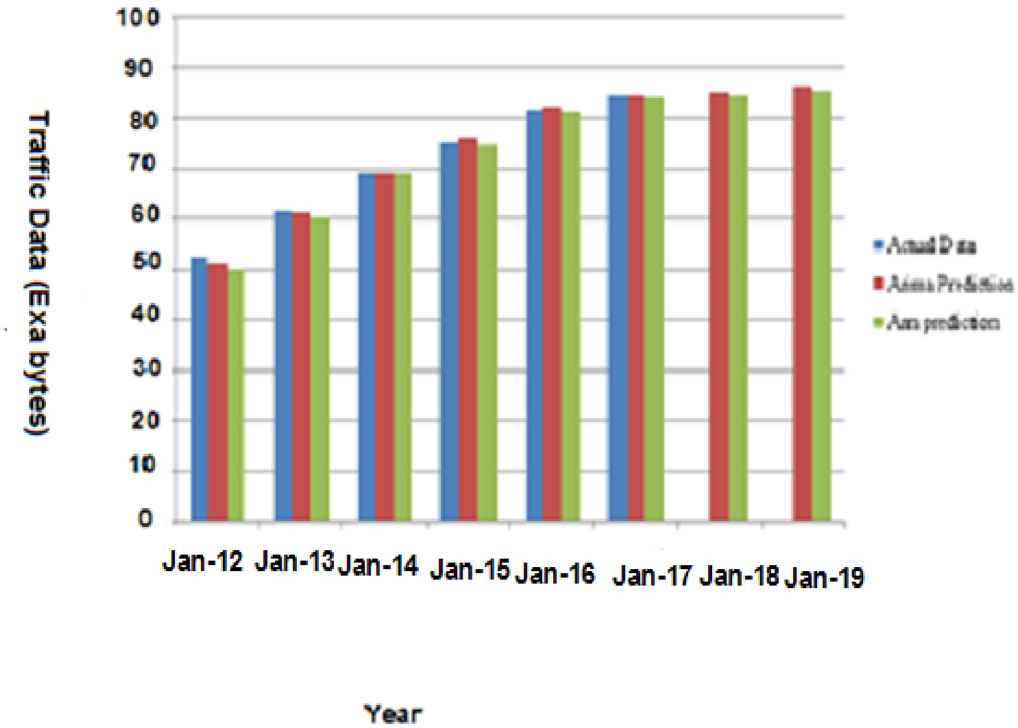

FIGURE 4-8: Evaluation between actual data, ARIMA and ANN

In the following section we compare the two methods based on their performance. We check if each method predicted accurately compared to the other. Table 4.5 shows a comparison of the feedback technique.

TABLE 4.5.Comparative study of feedback techniques

	Accurate Prediction	Future Prediction	Current Prediction	Faulty Prediction
ARIMA	Very High	High	High	Low
ANN	High	Medium	High	Medium

4.10. Comparison of prediction errors between ARIMA and ANN

To measure the accuracy of the algorithms and to compare which method has the least errors we will evaluate both methods by using Mean Absolute Percentage Error (MAPE). This tool is used to measure how close a fitted line is to data points (Almeida et al., 2015). The following section presents the overall measurements of several experimental simulations for determining the accurate traffic flow when using the Auto-Regressive Integrated Moving Average (ARIMA) model and the Artificial Neural Networks (ANNs) model. To evaluate the results between the two methods we apply the evaluation measures mentioned in section 3.4. The following is the summary of the results.

Figure 4-10 shows the comparison of prediction errors between the Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs) models. We can see that our models decreases prediction errors significantly for a long a time before the prediction of the future traffic and ARIMA outperforms ANN at some stage with regards to the overall prediction. The problem of the Auto-Regressive Integrated Moving Average (ARIMA) starts when it must predict high volume of data, where the Artificial Neural Networks (ANNs) model is very comfortable with any data. In the years 2012, 2014 and 2015 Artificial Neural Networks (ANNs) have more errors compared to the ARIMA model but gain confidence in the year 2017. Figure 4-9 shows the comparison of prediction errors between ARIMA and ANN by using Mean Absolute Percentage Error (MAPE) from 2011 to 2017.

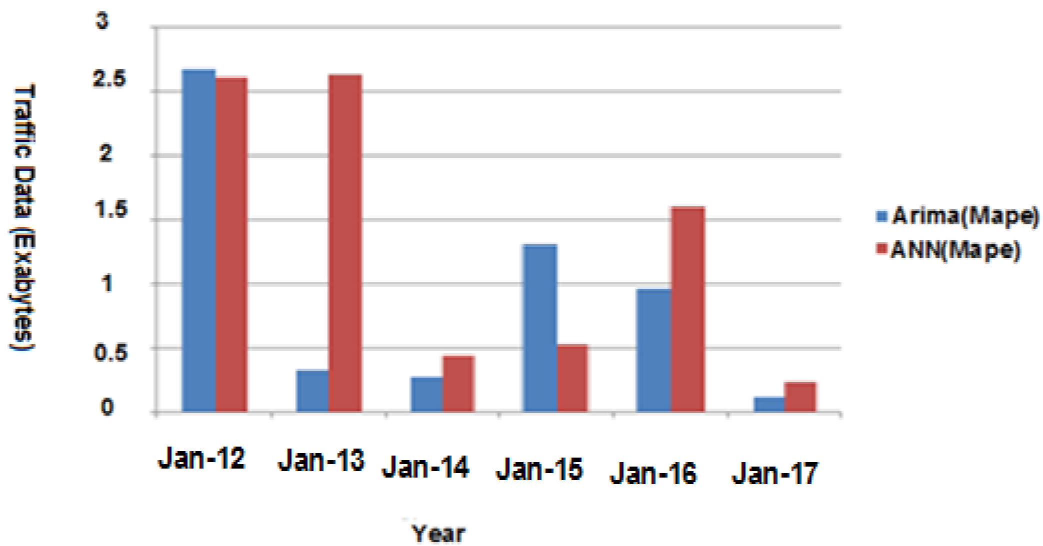

FIGURE 4-9: Errors calculated prediction in Figure(4-8)

As per Figure (4-9) of errors calculation Auto-Regressive Integrated Moving Average (ARIMA) has an average of 43 % of errors per year compared to the Artificial Neural Networks (ANNs) model which has 57 % of errors per year in this regard. Auto-Regressive Integrated Moving Average (ARIMA) minimise errors when it is performed on large volume of data in Exabyte. In Table 4-2, where Auto-Regressive Integrated Moving Average (ARIMA) was predicted per month in Exabyte, the MAPE error shows 1.15 % of prediction errors and in Table 4-4 where Artificial Neural Networks (ANNs) was predicted per month, the MAPE error was 0.325 % .We can conclude by saying that Artificial Neural Networks (ANNs) outperformed Auto-Regressive Integrated Moving Average (ARIMA) when performing with minimum data per month.

4.11. Discussion of the predicted results in cloud computing:

Figures 4-3 and 4-7 shows accurate traffic prediction in cloud computing, when compared to models used by other scholars the following improvements were accomplished:

- Elimination of prediction error frequency results due to space limits from the model;
- Accurate prediction from actual data;
- Elimination of faulty predictions and detect false alarms from the model;
- Maximized accurate future prediction;
- Elimination of admission control and bottleneck management from the model; and
- Effective congestion control.

4.12. Overall Results Analysis

The proposed method for this study, which is the ARIMA and ANN models, have been implemented and evaluated in detail. It has shown that traffic prediction can minimise prediction errors according to the requirements of the customers. The most challenging thing about traffic flow is that it is growing every day and more technology is needed to be able to accommodate the future volume of traffic in cloud computing.

When comparing this model to the models of other scholars the following improvements were accomplished: elimination of prediction error frequency results due to space limits from the model Gong et al., (2010), elimination of faulty predictions and detection of false alarms from the model of Gong et al., (2010). This Faulty prediction causes the system to be inaccurate. Elimination admission control and bottleneck management from the model of Song et al., (2012) for the allocation of resources.

For future prediction both Song et al., (2012) and Wan et al., (2014) do not show sizes of data centers and the volume of traffic to be predicted. This makes it impossible to do future bottleneck management and traffic prediction. From Figure 4-3 it can be observed that the two chosen models can predict traffic very well.

4.13. Summary of chapter 4

This chapter began with the process of designing a Auto-Regressive Integrated Moving Average (ARIMA) model where we learned that ARIMA has four (4) step before it can execute accurate results. Those steps are identification of the module, estimation of parameters, validations of the module and prediction. This was followed by the designing of a Artificial Neural Networks (ANNs) model where we learned that this technique has five (5) steps before it can produce the results, namely, input signal synaptic weights and summing junction followed by the activation function and then it will produce the output. After the design we simulated and implemented ARIMA by using Matlab. This method was used to predict traffic flow per year in Exabyte from 2011 to 2018 and the method predicted traffic very well. We then implemented the model per month for the year 2013.

The ARIMA model showed minimum errors on a monthly basis and the next step of our experiment was to simulate and implement ANN by using Matlab .This method was used to predict the traffic flow per year in Exabyte from 2012 to 2018 and we saw a decline in reducing errors per year. We then implemented the module per month for the year 2013 and we observed that ANN was reducing errors. In this chapter we also focused on the evaluation of errors between ARIMA and ANN. We observed that ARIMA outperformed ANN only when predicting per year. In this regard ARIMA can do very well when predicting the large volume of data.

Furthermore, we observed that ANN outperformed ARIMA when predicting on a monthly basis. In this regard Information Technology organization can implement ARIMA only for large volumes of data and implement ANN for a small volume of data. In this study we focused only on the prediction of traffic flow in cloud computing for data that was observed by Information Technology cloud organizations (Columbus, 2017). In the future there is a need to predict the resources that will be required for the client at that point in time.

CHAPTER 5

5. Conclusion, Contribution and Future Work

This chapter gives an deliberated conclusion, contribution and recommendations for future research. In the first chapter, one of our objectives was to "utilize accurate traffic flow prediction in cloud computing". There is a need to evaluate whether our objectives were achieved or not at the same time evaluating the contribution of the research work and make recommendations regarding future research.

5.1. Conclusion

Cloud computing provide clients with a network of remote servers hosted on the internet, to store, manage, and process data, rather than them having to own and manage a Lan server or a personal computer (e.g., networks, servers, storage, applications and services).

With this array of opportunities come some challenges amongst which are data, inconsistency of traffic flow, security threats, performance unpredictability and ensuring prompt (quick) resource scaling.

This thesis focused on predicting traffic accurately according to the requirements of the customer. Several techniques have existed in the area of prediction however accurate prediction was ignored. This study predicted traffic accurately according to the demand of users, looking and issues such as accurate traffic prediction, analysis of traffic flow and future traffic prediction.

Detailed discussions and challenges of cloud computing was discussed in chapter 2. Techniques for overcoming some challenges of accurate prediction where provided and presented in chapter 3, and the results of the models used was discussed in chapter 4. However, there are future research required that must be investigated to continue to reduce errors in traffic forecast.

5.2. The Main Contributions of this study are as Follows:

- Comprehensive reviews of subdivision (Traffic flow in cloud), representation and similar techniques were done, challenges, open issues, advantages and disadvantages were highlighted. Standard evaluations of these techniques were recommended for easy comparison. Research papers for international conferences were written.
- Analysis of traffic flow simulation using Matlab tool was demonstrated.
- The use of Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) techniques were proposed and a comprehensive evaluation was done. The techniques are suitable for accurate traffic prediction. The study contributed in accurate traffic prediction in cloud service and service providers to meet service level agreement.

5.3. Future Work

Traffic forecast in cloud computing is a very interesting area that is still under investigation because the cloud traffic is increasing every year. Traffic flow is inconsistent and it increases or even decreases every second or minute per day. Cloud computing technology for the traffic flow in cloud computing is still a challenging issue that do not provide the users of cloud services with the required services. There is a need to address future management of traffic and the allocation of resources.

In this review, traffic forecast by using the ARIMA and ANN methods has clarified trusting results. Be that as it may, a few areas have been isolated for further research and they are:

- This study has focused only on the traffic prediction. Further work to include the cost of using services, resource prediction and security in cloud computing may be worth investigating.

- Auto-Regressive Integrated Moving Average (ARIMA) must be further investigated for poor prediction when the model is predicting small volumes of Data.
- Artificial Neural Networks (ANN) is giving poor results when the model is predicting large volumes of data.
- The study focused on accurate traffic prediction and reduce faulty prediction. Further work is required to investigate how many servers and system requirements to connect and host this solution.
- The greatest challenge is getting 100% accurate prediction results compared to actual data. Therefore, there is still a need for further investigation in the area of accurate traffic prediction in cloud computing Systems.
- Investigating the use of a blend of ARIMA-ANN and determining what would happen later on system that may additionally build the announcement on what would happen later if there would be no errors.

5.4. Summary of chapter 5

This chapter looked at the overall contribution of the study and it was observed that Auto-Regressive Integrated Moving Average (ARIMA) has a limitation when predicting small volumes of data where Artificial Neural Networks (ANN) has a limitation when predicting large volumes of data. This study proposed the combination of ARIMA-ANN for consideration in future research. Furthermore we proposed the prediction of resources. In this study it was observed that it is possible to predict traffic flow and minimise errors according to the need of the customers or users. This study also showed that there is a challenge of getting 100% accurate prediction. Therefore there is a need for further investigation in the area of data traffic prediction.

REFERENCES

ALMEIDA, M. P., PERPIÑÁN, O. & NARVARTE, L. 2015. PV power forecast using a nonparametric PV model. *Solar Energy,* 115**,** 354-368.

BELOGLAZOV, A. & BUYYA, R. 2012. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in clouddata centers. *Concurrency and Computation: Practice and Experience,* 24**,** 1397-1420.

BLUME, O., ECKHARDT, H., KLEIN, S., KUEHN, E. & WAJDA, W. M. 2010. Energy savings in mobile networks based on adaptation to traffic statistics. *Bell Labs Technical Journal,* 15**,** 77-94.

BOTO-GIRALDA, D., DÍAZ-PERNAS, F. J., GONZÁLEZ-ORTEGA, D., DÍEZ-HIGUERA, J. F., ANTÓN-RODRÍGUEZ, M., MARTÍNEZ-ZARZUELA, M. & TORRE-DÍEZ, I. 2010. Wavelet-Based Denoising for Traffic Volume Time Series Forecasting with Self-Organizing Neural Networks. *Computer-Aided Civil and Infrastructure Engineering,* 25**,** 530-545.

BOYD, S., PARIKH, N., CHU, E., PELEATO, B. & ECKSTEIN, J. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning,* 3**,** 1-122.

BUYYA, R., RANJAN, R. & CALHEIROS, R. N. 2010. Intercloud: Utility-oriented federation of CLOUD COMPUTING environments for scaling of application services. *Algorithms and architectures for parallel processing.* Springer.

CHANG, N.-B., BAI, K. & CHEN, C.-F. 2015. Smart information reconstruction via time-space-spectrum continuum for cloudremoval in satellite images. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of,* 8**,** 1898-1912.

DALMAZO, B. L., VILELA, J. P. & CURADO, M. Online Traffic Prediction in the Cloud: A Dynamic Window Approach. Future Internet of Things and Cloud(FiCloud), 2014 International Conference on, 2014. IEEE, 9-14.

DERI, L. & FUSCO, F. MicroCloud-based network traffic monitoring. Integrated Network Management (IM 2013), 2013 IFIP/IEEE International Symposium on, 2013. IEEE, 864-867.

DI, S., KONDO, D. & CIRNE, W. Host load prediction in a Google compute cloudwith a Bayesian model. Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, 2012. IEEE Computer Society Press, 21.

DINH, H. T., LEE, C., NIYATO, D. & WANG, P. 2013. A survey of mobile CLOUD COMPUTING : architecture, applications, and approaches. *Wireless communications and mobile computing,* 13**,** 1587-1611.

DUY, T. V. T., SATO, Y. & INOGUCHI, Y. Performance evaluation of a green scheduling algorithm for energy savings in CLOUD COMPUTING . Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW), 2010 IEEE International Symposium on, 2010. IEEE, 1-8.

GANAPATHI, A., CHEN, Y., FOX, A., KATZ, R. & PATTERSON, D. Statistics-driven workload modeling for the cloud. Data Engineering Workshops (ICDEW), 2010 IEEE 26th International Conference on, 2010. IEEE, 87-92.

GANAPATHI, A., KUNO, H., DAYAL, U., WIENER, J. L., FOX, A., JORDAN, M. & PATTERSON, D. Predicting multiple metrics for queries: Better decisions enabled by machine learning. Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on, 2009. IEEE, 592-603.

COLUMBUS, L. 2016. COLUMBUS, L. 2016. cloud-computing-forecasts A Passion for Research,Focusing on CRM, CLOUD COMPUTING , ERP and enterprise software. CLOUD COMPUTING Index.

GILL, P., JAIN, N. & NAGAPPAN, N. Understanding network failures in data centers: measurement, analysis, and implications. ACM SIGCOMM Computer Communication Review, 2011. ACM, 350-361.

GONG, Z., GU, X. & WILKES, J. Press: Predictive elastic resource scaling for cloudsystems. Network and Service Management (CNSM), 2010 International Conference on, 2010. IEEE, 9-16.

GOVINDAN, S., LIU, J., KANSAL, A. & SIVASUBRAMANIAM, A. Cuanta: quantifying effects of shared on-chip resource interference for consolidated virtual machines. Proceedings of the 2nd ACM Symposium on CLOUD COMPUTING , 2011. ACM, 22.

GROBAUER, B., WALLOSCHEK, T. & STÖCKER, E. 2011. Understanding CLOUD COMPUTING vulnerabilities. *Security & privacy, IEEE,* 9**,** 50-57.

GROSSMAN, R. & WHITE, K. 2012. A vision for a biomedical cloud. *Journal of internal medicine,* 271**,** 122-130.

HERBST, N. R., HUBER, N., KOUNEV, S. & AMREHN, E. 2014. Self-adaptive workload classification and forecasting for proactive resource provisioning. *Concurrency and Computation: Practice and Experience,* 26**,** 2053-2078.

HU, J., GU, J., SUN, G. & ZHAO, T. A scheduling strategy on load balancing of virtual machine resources in CLOUD COMPUTING environment. Parallel Architectures, Algorithms and Programming (PAAP), 2010 Third International Symposium on, 2010. IEEE, 89-96.

IQBAL, W., DAILEY, M. N., CARRERA, D. & JANECEK, P. 2011. Adaptive resource provisioning for read intensive multi-tier applications in the cloud. *Future Generation Computer Systems,* 27**,** 871-879.

ISHAK, S. & AL-DEEK, H. 2002. Performance evaluation of short-term time-series traffic prediction model. *Journal of Transportation Engineering,* 128**,** 490-498.

ISLAM, S., KEUNG, J., LEE, K. & LIU, A. 2012. Empirical prediction models for adaptive resource provisioning in the cloud. *Future Generation Computer Systems,* 28**,** 155-162.

ISRAIL, S., PANDEY, R. & CHOURASIA, U. 2015. Dynamic Resource Allocation for Overload Avoidance and Green CLOUD COMPUTING . *International Journal of Computer Applications,* 126.

KALBASI, A., KRISHNAMURTHY, D., ROLIA, J. & DAWSON, S. 2012. Dec: Service demand estimation with confidence. *Software Engineering, IEEE Transactions on,* 38**,** 561-578.

KHASHEI, M., BIJARI, M. & ARDALI, G. A. R. 2009. Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs). *Neurocomputing,* 72**,** 956-967.

KOUKI, Y., SABBIR HASAN, M. & LEDOUX, T. Delta Scaling: How Resources Scalability/Termination Can Be Taken Place Economically? Services (SERVICES), 2015 IEEE World Congress on, 2015. IEEE, 55-62.

LACURTS, K. L. 2014. *Application workload prediction and placement in CLOUD COMPUTING systems.* Massachusetts Institute of Technology.

LACURTS, K., MOGUL, J. C., BALAKRISHNAN, H. & TURNER, Y. 2014. Cicada: Predictive Guarantees for CloudNetwork Bandwidth.

LAGAR-CAVILLA, H. A., WHITNEY, J. A., SCANNELL, A. M., PATCHIN, P., RUMBLE, S. M., DE LARA, E., BRUDNO, M. & SATYANARAYANAN, M. SnowFlock: rapid virtual machine cloning for CLOUD COMPUTING . Proceedings of the 4th ACM European conference on Computer systems, 2009. ACM, 1-12.

LI, Z., ZHANG, M., ZHU, Z., CHEN, Y., GREENBERG, A. G. & WANG, Y.-M. WebProphet: Automating Performance Prediction for Web Services. NSDI, 2010. 143-158.

LIU, L., BAR, J., FRIEDRICH, F., GUTKNECHT, J. & TSAO, S.-L. A low power configurable SoC for simulating delay-based audio effects. Reconfigurable Computing and FPGAs (ReConFig), 2012 International Conference on, 2012. IEEE, 1-6.

LOPES DALMAZO, B., VILELA, J. P. & CURADO, M. Predicting Traffic in the Cloud: A Statistical Approach. Cloudand Green Computing (CGC), 2013 Third International Conference on, 2013. IEEE, 121-126.

MANVI, S. S. & SHYAM, G. K. 2014. Resource management for Infrastructure as a Service (IaaS) in CLOUD COMPUTING : A survey. *Journal of Network and Computer Applications,* 41**,** 424-440.

MARK, C. C. T., NIYATO, D. & CHEN-KHONG, T. Evolutionary optimal virtual machine placement and demand forecaster for CLOUD COMPUTING . Advanced Information Networking and Applications (AINA), 2011 IEEE International Conference on, 2011. IEEE, 348-355.

MARSTON, S., LI, Z., BANDYOPADHYAY, S., ZHANG, J. & GHALSASI, A. 2011. CLOUD COMPUTING —The business perspective. *Decision support systems,* 51**,** 176-189.

MELL, P. & GRANCE, T. 2011. The NIST definition of CLOUD COMPUTING .

MOAYEDI, H. Z. & MASNADI-SHIRAZI, M. ARIMA model for network traffic prediction and anomaly detection. Information Technology, 2008. ITSim 2008. International Symposium on, 2008. IEEE, 1-6.

NGUYEN, H., SHEN, Z., TAN, Y. & GU, X. FChain: Toward black-box online fault localization for cloudsystems. Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference on, 2013. IEEE, 21-30.

NGUYEN, T. T., ARMITAGE, G., BRANCH, P. & ZANDER, S. 2012. Timely and continuous machine-learning-based classification for interactive IP traffic. *IEEE/ACM Transactions on Networking (TON),* 20**,** 1880-1894.

NIU, D., XU, H., LI, B. & ZHAO, S. Quality-assured cloudbandwidth auto-scaling for video-on-demand applications. INFOCOM, 2012 Proceedings IEEE, 2012. IEEE, 460-468.

PENG, Y., CHEN, K., WANG, G., BAI, W., MA, Z. & GU, L. Hadoopwatch: A first step towards comprehensive traffic forecasting in CLOUD COMPUTING . INFOCOM, 2014 Proceedings IEEE, 2014. IEEE, 19-27.

PETKOVIĆ, D., ĆOJBAŠIĆ, Ž., NIKOLIĆ, V., SHAMSHIRBAND, S., KIAH, M. L. M., ANUAR, N. B. & WAHAB, A. W. A. 2014. Adaptive neuro-fuzzy maximal power extraction of wind turbine with continuously variable transmission. *Energy,* 64**,** 868-874.

POPA, L., YALAGANDULA, P., BANERJEE, S., MOGUL, J. C., TURNER, Y. & SANTOS, J. R. ElasticSwitch: practical work-conserving bandwidth guarantees for CLOUD COMPUTING . ACM SIGCOMM Computer Communication Review, 2013. ACM, 351-362.

P. Sekwatlakwatla, M. Mphahlele, T. Zuva, "Traffic flow prediction in cloud computing ", Proc. 18th Conf. Advances in Comp. and Comm. Engineering, pp. 123-128, 2016..

PREVOST, J. J., NAGOTHU, K., KELLEY, B. & JAMSHIDI, M. Prediction of clouddata center networks loads using stochastic and neural models. System of Systems Engineering (SoSE), 2011 6th International Conference on, 2011. IEEE, 276-281.

QIBO, S., SHANGGUANG, W., DANFENG, Y. & FANGCHUN, Y. ARM-CPD: detecting SYN flooding attack by traffic prediction. Broadband Network & Multimedia Technology, 2009. IC-BNMT'09. 2nd IEEE International Conference on, 2009. IEEE, 443-447.

RADHAKRISHNAN, S., PAN, R., VAHDAT, A. & VARGHESE, G. 2012. Netshare and stochastic netshare: predictable bandwidth allocation for data centers. *ACM SIGCOMM Computer Communication Review,* 42**,** 5-11.

REN, X., LIN, R. & ZOU, H. A dynamic load balancing strategy for CLOUD COMPUTING platform based on exponential smoothing forecast. CLOUD COMPUTING and Intelligence Systems (CCIS), 2011 IEEE International Conference on, 2011. IEEE, 220-224.

ROY, N., DUBEY, A. & GOKHALE, A. Efficient autoscaling in the cloudusing predictive models for workload forecasting. CLOUD COMPUTING (CLOUD), 2011 IEEE International Conference on, 2011. IEEE, 500-507.

SADEGHI, A.-R., SCHNEIDER, T. & WINANDY, M. 2010. Token-based CLOUD COMPUTING . *Trust and Trustworthy Computing.* Springer.

SAMANTA, B. & AL-BALUSHI, K. 2003. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical systems and signal processing,* 17**,** 317-328.

SARIPALLI, P., KIRAN, G., SHANKAR, R. R., NARWARE, H. & BINDAL, N. Load prediction and hot spot detection models for autonomic CLOUD COMPUTING . Utility and CLOUD COMPUTING (UCC), 2011 Fourth IEEE International Conference on, 2011. IEEE, 397-402.

SHAH, M. M. D., KARIYANI, M. A. A. & AGRAWAL, M. D. L. 2013. Allocation Of Virtual Machines In CLOUD COMPUTING Using Load Balancing Algorithm. *International Journal of Computer Science and Information Technology & Security (lJCSlTS), ISSN***,** 2249-2256.

SMITH, B. L., WILLIAMS, B. M. & OSWALD, R. K. 2002. Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies,* 10**,** 303-321.

SONG, D., SHI, E., FISCHER, I. & SHANKAR, U. 2012. Clouddata protection for the masses. *Computer***,** 39-45.

SU, H. & YU, S. 2007. Hybrid GA based online support vector machine model for short-term traffic flow forecasting. *Advanced Parallel Processing Technologies.* Springer.

SUTHAHARAN, S. 2014. Big data classification: Problems and challenges in network intrusion prediction with machine learning. *ACM SIGMETRICS Performance Evaluation Review,* 41**,** 70-73.

TAMMARO, D., DOUMITH, E. A., ZAHR, S. A., SMETS, J.-P. & GAGNAIRE, M. Dynamic resource allocation in cloudenvironment under time-variant job requests. CLOUD COMPUTING Technology and Science (CloudCom), 2011 IEEE Third International Conference on, 2011. IEEE, 592-598.

TAMMARO, D., VALENTI, S., ROSSI, D. & PESCAPÉ, A. 2012. Exploiting packet-sampling measurements for traffic characterization and classification. *International Journal of Network Management*, 22**,** 451-476.

TODE, H., NISHII, K. & TANIGAWA, Y. A participating fine-granular CLOUD COMPUTING platform with in-network guidance. Communications (ICC), 2014 IEEE International Conference on, 2014. IEEE, 4023-4028.

TOMÁS, L. & TORDSSON, J. Improving cloudinfrastructure utilization through overbooking. Proceedings of the 2013 ACM Cloudand Autonomic Computing Conference, 2013. ACM, 5.

VLAHOGIANNI, E. I., KARLAFTIS, M. G. & GOLIAS, J. C. 2014. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, 43**,** 3-19.

WAN, J., ZHANG, D., SUN, Y., LIN, K., ZOU, C. & CAI, H. 2014. VCMIA: a novel architecture for integrating vehicular cyber-physical systems and mobile CLOUD COMPUTING . *Mobile Networks and Applications*, 19**,** 153-160.

WANG, C.-F., HUNG, W.-Y. & YANG, C.-S. A prediction based energy conserving resources allocation scheme for CLOUD COMPUTING . Granular Computing (GrC), 2014 IEEE International Conference on, 2014. IEEE, 320-324.

WANG, C., CAO, N., REN, K. & LOU, W. 2012a. Enabling secure and efficient ranked keyword search over outsourced clouddata. *Parallel and Distributed Systems, IEEE Transactions on*, 23**,** 1467-1479.

WANG, G., NG, T. & SHAIKH, A. Programming your network at run-time for big data applications. Proceedings of the first workshop on Hot topics in software defined networks, 2012b. ACM, 103-108.

XU, H., MICHAEL, Y. S., WONG, E. Y., UPHOFF, J., XU, Y., TREADWAY, J. A., TRUONG, A., O'BRIEN, E., ASQUITH, S. & STUBBINS, M. 2003. Multiplexed SNP genotyping using the Qbead™ system: a quantum dot-encoded microsphere-based assay. *Nucleic Acids Research*, 31**,** e43-e43.

YE, Z., ZHOU, X. & BOUGUETTAYA, A. Genetic algorithm based QoS-aware service compositions in CLOUD COMPUTING . Database systems for advanced applications, 2011. Springer, 321-334.

YINGYU, Z. X. Z. L. A. & YINGXU, S. M. W. 2008. Analysis on the Structure and Characteristic of Weak Hail CloudRadar Echo [J]. *Meteorological Monthly,* 2**,** 007.

YOON, M. S., KAMAL, A. E. & ZHU, Z. 2015. Request Prediction in Cloudwith a Cyclic Window Learning Algorithm. *arXiv preprint arXiv:1507.02372*.

ZANDER, J. & MAHONEN, P. 2013. Riding the data tsunami in the cloud: myths and challenges in future wireless access. *Communications Magazine, IEEE,* 51**,** 145-151.

ZHANG, Q., CHENG, L. & BOUTABA, R. 2010. CLOUD COMPUTING : state-of-the-art and research challenges. *Journal of internet services and applications,* 1**,** 7-18.

ZHANG, Y., ROUGHAN, M., LUND, C. & DONOHO, D. An information-theoretic approach to traffic matrix estimation. Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications, 2003. ACM, 301-312.

ZOHAR, E., CIDON, I. & MOKRYN, O. 2014. PACK: Prediction-based cloudbandwidth and cost reduction system. *IEEE/ACM Transactions on Networking (TON),* 22**,** 39-51.

ZOHAR, E., CIDON, I. & MOKRYN, O. O. The power of prediction: Cloudbandwidth and cost reduction. ACM SIGCOMM Computer Communication Review, 2011. ACM, 86-97.