Journal of Computer Science

Original Research Paper

Computational and Statistical Analysis of Security and Privacy Parameters of Cloud Computing in Information Technology

[1*] Jimbo Claver, [2] Edris Hamraz, [3] Takeru Suzuki, [4] Ngongo Seraphin, [5] Andjiga Gabriel and [6] Etoua Magloire

[1] *Department of Applied Mathematics, MMAFSP, Rise Waseda University, Tokyo, Japan*
[2] *Department of Information Technology, American University of Afghanistan, Kabul, Afghanistan*
[3] *Department of Applied Mathematics and Statistics, Rise Waseda University, Tokyo, Japan*
[4] *Department of Applied Mathematics, MMAFS -ENSY, University of Yaounde-1, Yaoundé, Cameroon*
[5] *Department of Applied Mathematics, MMAFSP, University of Yaounde-1, Yaoundé, Cameroon*
[6] *Department of Applied Mathematics, MMAFSP, Advanced School of Engineering of Yaoundé (NASEY), Cameroon*

Article history
Received: 11-07-2019
Revised: 14-11-2020
Accepted: 23-11-2020

Corresponding Author:
Jimbo Claver
Department of Applied Mathematics, MMAFSP, Rise Waseda University, Tokyo, Japan
Email: jimbo.maths@gmail.com

Abstract: Information Technology has transformed the way cloud computing itself can help to manage, consume and improve cost efficiencies, accelerate innovation, provide faster time to the market and develop the ability to scale applications on demands and services. In this research work, we present recent developments and some statistical analysis techniques used in cloud computing development with challenges related to the security and privacy in information technology. The ultimate goal in this research is to detect relevant factors that are more likely to affect security and privacy in cloud computing services. Moreover, we identify the security risks and the threats that have already been recognized by cloud security alliance as vital components of those challenges. Finally, the solutions we found in this study will efficiently help private and governmental organizations to resolve some of the important security challenges. Although the suggested solutions still remain at the Service level agreements contracted between the cloud provider and consumer, it will secure the cloud users and earn their satisfactions in cloud computing services.

Keywords: Cloud Computing, Security, Privacy, Statistical Data Analysis, Parameters, Information Science, Statistical Modelling and Data Mining

Introduction

Recently, most of the organizations, service providers and enterprises are interested in implementing clouds to remove the challenge of integration in complex software and hardware components. Cloud computing platforms are more attractive because they let a business quickly moved and able to host private and public resources on demand without the complexities and time association regarding the purchase, installation, configuration and deployment of traditional physical infrastructure. (Kandukuri *et al*., 2009) and (Holstein and Stouffer, 2010). The goal of this study is to analyze the security of consumer and security of cloud provider with their impact on the organizations that use cloud computing services. When the companies want to use the latest technology like cloud computing services, they will find out security and privacy as well as some other challenges (Ristenpart *et al*., 2009; Grobauer *et al*., 2010; Zhang, 2010). The main contribution of this study is to provide some foundations about quantitative research on issues, challenges of privacy and security of cloud computing. As such, this research represents the first in its kind practical and viable solutions to cloud computing problems are explored. We mainly rely on Service Level Agreements (SLA) between cloud service provider and customers and also on some technical analysis. In addition, we develop strategies for capturing the risks

Science Publications

© 2020 Jimbo Claver, Edris Hamraz, Takeru Suzuki, Jawad Azimi, Andjiga Gabriel and Etoua Magloire. This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license.

associated to solutions to these issues in Information Technology (Microsoft, 2010).

Background Research

Cloud Computing

Considering that cloud computing is still a relatively new innovation to the world of technology, a lot of avenues for quantitative and technological researches have remained open. There are various open issues that need to be resolved before cloud computing is fully accepted by the broad community. Before we will dive into the research methodology of this research work, a deeper explanation is needed on what cloud computing is in reality. Previously, we provide the initial definition of cloud computing and now we will extend this definition to cloud computing models relying respectively on the work of the following authors: (Jimbo, 2019; Claver *et al*., 2018; Subashini and Kavitha, 2011; OWASP, 2010; Popović and Hocenski, 2010; Peter and Tim, 2009; Microsoft, 2010; Bozdogan, 2000).

As state in (Jensen *et al*., 2009; ENISA, 2010) "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction". This definition is supported by five key cloud characteristics, three delivery models and four deployment models. These supporting properties will be explained in the next section, we will also discuss various security issues and concerns related to the structure of cloud computing itself. The NIST cloud definition framework includes cloud definition and models discussed as following in Fig. 1.

Deployment Models

Hybrid Cloud

Hybrid cloud refers to a mixed computing, storage and services environment made up of on-premises infrastructure, private cloud services and a public cloud-such as Amazon Web Services (AWS) or Microsoft Azure-with orchestration among the various platforms.

Public Cloud

A public cloud is a type of computing in which a service provider makes resources available to the public via the internet. Resources vary by provider but may include storage capabilities, applications or virtual machines. Public cloud allows for scalability and resource sharing that would not otherwise be possible for a single organization to achieve.

Cloud Service Models

Software-as-a-Service (SaaS)

This model provides service called Software-as-a-Service (SaaS). The SaaS service provides services as an application to the consumer, with standard interfaces. These services will not be visible to the consumer and works on the background of services at the top of the infrastructure. The responsibility goes straight to the cloud provider and those services are application management, operating system and basic of infrastructure (CSA, 2010; Mell and Grance, 2011),

Platform-as-a-Service (PaaS)

PaaS service model offers a model for operation as services and development of platforms to the consumer. "The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations" (Mell and Grance, 2011; Elizabeth and Vadim, 2007).

Infrastructure-as-a-Service (IaaS)

Infrastructure as services is the lowest service model in the technology stack and this service offered services such as raw data storage, processing power and network capacity. The consumer can use IaaS based service offerings to deploy his own operating systems and applications, offering a wider variety of deployment possibilities for a consumer than the PaaS and SaaS models. "The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications and possibly limited control of select networking components (e.g., host firewalls)" (Mell and Grance, 2011).

The Architecture of Cloud Computing

They represent a top-level architecture of cloud that depicts various cloud service delivery models. Cloud enhanced collaboration, agility, scalability, availability and provides the potential for cost reduction through optimized and efficient computing. From an architectural perspective, given this abstracted evolution of technology, there is much confusion surrounding how the cloud is both similar and different from existing models and how these similarities and differences might impact the organizational, operational and technological approaches to cloud adoption as it relates to traditional network and information security practices.

DOI: 10.3844/jcssp.2020.1625.1638

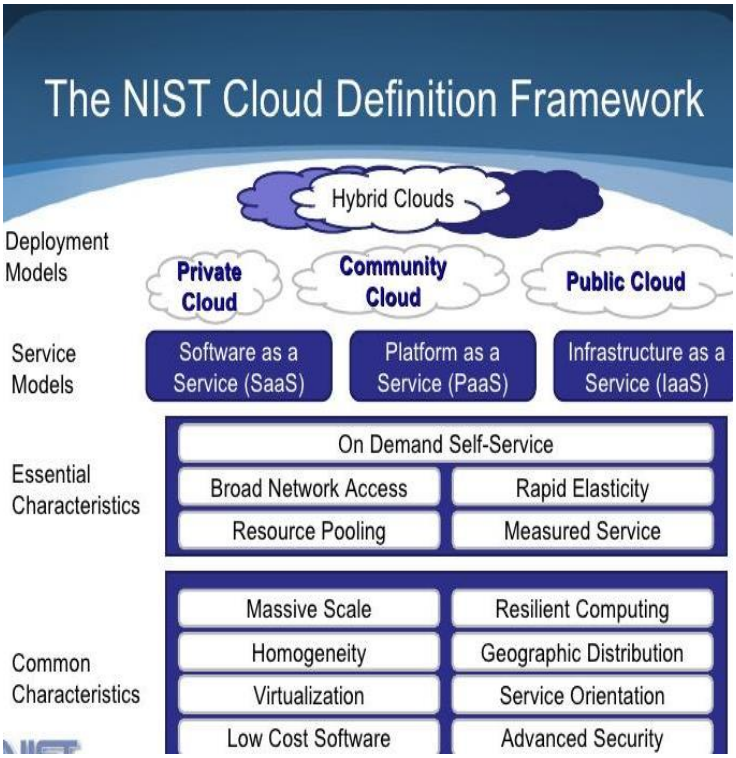


Fig. 1: Cloud computing definition (NIST, 2010)

Security and Privacy Challenges in Cloud Computing

In the introduction, we discuss the technology of the cloud and the importance of Security in the cloud. This article illustrates the unique issues of cloud computing that show security and privacy challenges in clouds and also discusses various approaches to address these challenges and explores the future work needed to provide a trustworthy cloud computing environment. The unique security and privacy implications of cloud computing rely on outsourcing data and applications, service level agreement, heterogeneity, compliance regulations, virtualization and hypervisors, authentication and identity management, trust management and policy integration, access control and accounting, service-service management, privacy and data protection.

Although security and privacy services in the cloud can be fine-tuned and managed by experienced groups that can potentially provide efficient security management and threat assessment services, the issues we've discussed here show that existing security and privacy solutions must be critically reevaluated with regard to their appropriateness for clouds. Many enhancements in existing solutions, as well as more mature and newer solutions, are urgently needed to ensure that cloud computing benefits are fully realized as its adoption accelerates. Cloud computing is still in its infancy and how the security and privacy landscape changes will impact its successful, widespread adoption (Chen *et al*., 2012)

Classification of Security Issues in Cloud Computing

The security issues in cloud computing can be categorized into the following three broad classes:

- *Traditional security concerns*
- *Availability issues*
- *Third party data control-related issues*

Security in the cloud is one the factors by which more organizations effect and lost their confidentiality. Due to the security and privacy risks, most of the companies are not willing to adopt the cloud computing services; they have the following concerns.

Traditional Security Concern

The companies are thinking when the have used the cloud computing services automatically lost control of their own company.

Availability Issues

The companies think when something happened to the cloud provider they will lose everything but this is not correct because the company are investing more than millions of dollars into reliability of data.

Third Party Control

The companies think when we put our data to the cloud we will lost control of everything and confidentiality of the data but it is not correct because the current system allows the confidentiality of the data.

Traditional security concern is one of the factors in cloud computing and those factors can impact the security of cloud computing. Besides, most of companies accept cloud for availability and cloud computing gives the organization 99.99% availability, it is too costly for organization to have this kind of support for the users. There is the third party control that most of the organizations fear of adopting cloud computing and they think while they accept cloud computing their data will be leaked and that can impact the cloud computing services in many important organizations.

The security issues involve computer and network intrusions or attacks that will be made possible or at least easier by moving to the cloud. Cloud providers respond to these concerns by arguing that their security measures and processes are more mature and tested than those of the average companies. Another argument made by the Jericho Forum (don't cloud vision) is "It could be easier to lock down information if it's administered by a third party rather than in-house if companies are worried about insider threats… In addition, it may be easier to enforce security via contracts with online services providers than via internal controls. Many other concerns access and managed by internal controls as well.

Research Methodology

This section will provide information about how the project will be covered in order to obtain the desired results. We will provide information about how cloud computing works, how it will be integrated with security and privacy issues of cloud computing that have been used in their organization and we will find out security and privacy issues by using case study method to reach our research goals. For conducting this research, we will design some questionnaires and send them through emails to all organizations in Afghanistan and also conduct web research and literature reading to find appropriate questionnaires to the problem.

Method and Process

We have used case study method in this study because there are many research papers that already exist and they have found out many issues related to cloud computing. We mainly focus on Afghanistan because of the availability of the data and also the level of trust we can have. We decided to spread the questionnaire to many organizations in Afghanistan to understand the problems they face and provide them with proper solutions. The ultimate goal in this research work is to analyzed the impact of the cloud computing security and privacy issues in overall public and private organizations, especially in organizations that somehow are using the services of cloud computing. Their survey is very important to us and the result of this survey depend on the results of the questionnaire that have been sends trough the email. The analysis will help organizations to find the best cloud provider and select which point they should choose their cloud provider and what are the important security and issues to be consider when adopting cloud computing services.

The easiest method was specifically developed to find the security and privacy concerns that most organizations had regarding their protection and safety. The important point was that cloud provider wanted to ignore such concerns and most organizations in Afghanistan were not aware of such thing like security and privacy issues and SLA Service Level Agreement that will be used to solve the problems in this research work.

Data Collection

Source of Data Collection

We used the direct mailing procedure to send the prepared questionnaire to the experts and few of responses are taken from direct interaction with the experts who are currently having industrial experience in cloud computing. The questionnaire was sent to many governments and International NGO to reach out their comments regarding security and privacy issues of cloud computing. Also, we use the interview component to gather additional information.

Survey Questions Formation

In this process, we designed the questionnaire to answer the research questions. The process involves that we designed questionnaire and discussed with colleagues to make the questionnaire error free and correct that ensures the designed questionnaire truly reflects the aims and objectives of the research. During this process, we designed the questionnaire according to problem statements and how to find those problems that exist in the organization of Afghanistan. In the questionnaire, we include some General question about the level of knowledge department and organization they are working on it alongside years of experience in it industry.

Survey Administration

In this research, we adopt the method of electronic survey. This process includes the identification of experts in cloud computing with relevant experience in cloud computing and sent the requesting mail which contains the brief introduction of us, topic area and purpose of conducting the survey and requesting to participate in the survey. After the acknowledgment from the experts, we sent them a questionnaire to answer them. This survey is conducted in December of 2016 until May 2017. It has taken a long time to get a number of responses because of unavailability of experts. The questionnaire starts with the name of the expert, organization name, experience in cloud computing, email, contact number and the year of experience in IT industry. The detailed description about the survey participants is explained in the results.

Survey Results on Google Form

The link for google form survey is below, that we have been contacted much organizations in Afghanistan, most of them working in IT industry and or involved in this technology. The goal was to find who is interested in cloud computing and why there is fear of the technology. But unfortunately almost 80% of customers have feared to use more services of cloud computing because of security and privacy issues. In the figure below, the reader will be able to see the results of the people interviewed through email addresses with complete information that was needed.

As per the survey that we have done from international and governmental organizations in Afghanistan we have found that many organizations in Afghanistan are not using cloud computing services due to limitations and lack of awarenesses of cloud computing services. For them, security and privacy issues represent a big challenge for the organizations, especially in Afghanistan, most of the organizations do not have experts or professional advisors to guide the companies in order to be aware of the advanced features of cloud computing and to use them in the same time.

Alongside the benefits that cloud computing has in Afghanistan, all of the organizations are still reluctant for using this new technology. When researches are conducted, many experts from IT governmental organizations in Afghanistan, private sectors, international NGO, some universities were also not opened to this new technology. These organizations are mostly using parts of cloud technology. In this study, we use the Google form questionnaires which were conducted by surveys.

Data Structure and Characteristics

The data have been collected from many international and governmental organizations in Afghanistan. The criteria for selecting these organizations were based on their awareness of cloud computing and most of them were using cloud computing services and few of them already experienced the effectiveness of security and privacy in cloud computing. The best way of collecting data from international and governmental organizations was through google form, we send the questionnaires electronically to many organizations and collected the accurate information from those organizations.

In the questionnaires, we gathered the information regarding the security and privacy of cloud computing in order to find the impact of security and privacy of these organizations. The questionnaires included: Questions regarding security risks, security policy, standards, encryption technologies, internet attacks and recovery procedures after an attack on the organization. In the next section, we will present the data modelling and analysis of this research work.

Statistical Modelling

Multiple Linear Regression

A multiple OLS Regression is used when the dependent variable is continuous and there are more than one independent variables. The OLS regression is used to predict the changes in the dependent variable based on the changes in the values of the independent variables. This is the estimated OLS multiple regression equation (Kutner *et al*., 2005):

$$y = b_0 + b_1 x_{i1} + b_2 x_{i2} + ... + b_k x_{ik} + e_i \quad (1)$$

where, y is the value of response variable, b_0 is estimated intercept, (b_1,b_2,+…..+b_k) are the slopes of regression line, (x_{i1}+x_{i2}+…..+x_{ik}) are the observations and e is the error term. The multiple linear regression is used to assess the relation between dependent variable and independent variables of the study. The five underlying assumptions of a multiple OLS regression are tested to make sure the OLS estimators are the Best Linear Unbiased Estimators (BLUE) of b_0 and b_i.

Linearity

In simple regressions, linearity is evaluated by drawing the scatter plot of the dependent and independent variables. However, in multiple regressions, we intuitively judge whether or not the variables have linear regression. Mostly, a high R-square indicates the variables are linearly related with each other.

Normality

One of the underlying assumptions of OLS Regression is normal distribution of the residuals. To evaluate the model for normality assumption, there are three ways. A histogram of the residuals can give an idea whether or not the residuals are normally distributed. However, to be more precise to decide and judge, Jarque-Bera, Skewness and Kurtosis tests are conducted. While Jarque-Bera test results help us to know about normality of the residuals, the later test gives information about Skewness and Kurtosis separately (Bozdogan, 2000). These tests calculate the p-value to reject or accept null hypothesis of normality. A *p-value* of 0.05 or low rejects the null hypothesis of normality; thus, the residuals are considered to be normally distributed when the p-value is not larger than common threshold value of 0.05. The Skewness and Kurtosis tests give two p-values one for Skewness and one for Kurtosis. Therefore, this test gives the necessary information to know whether the distribution of residuals have Skewness or Kurtosis or both problems. This is the formula for calculating Jarque-Bera test for normality in multiple regression analysis (Elizabeth and Vadim, 2007):

$$JB = \frac{n-k}{6}\left(S^2 + \frac{1}{4}(C-3)^2\right) \tag{2}$$

Where:
JB = Jarque-Bera test
N = Number of observations
S = Sample skewness
C = Sample kurtosis
K = Number of regressor (s)

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}_3} = \frac{\frac{1}{n}\sum_1^n \left(x_i = \hat{x}\right)^3}{\left(\frac{1}{n}\sum_1^n \left(x_i = \hat{x}\right)^2\right)^{3/2}} \tag{3}$$

$$C = \frac{\hat{\mu}_4}{\hat{\sigma}_4} = \frac{\frac{1}{n}\sum_1^n \left(x_i = \hat{x}\right)^4}{\left(\frac{1}{n}\sum_1^n \left(x_i = \hat{x}\right)^2\right)^{2}} \tag{4}$$

Where:
$\hat{\mu}_3$ = Estimates of third central moment
$\hat{\mu}_4$ = Estimates of fourth central moment
$\bar{x}$ = Sample mean
$\hat{\sigma}^2$ = Estimates of second central moment, variance

Heteroskedasticity

The easiest way to check for heteroscedasticity is drawing the scatter plot of residuals against the dependent variables of fitted plots. However, it is sometimes difficult to make a decision confidently about existence of heteroscedasticity based on such scatter plot. When deciding about heteroscedasticity, it is difficult based on the information given by the scatter plot, researchers use (Bin *et al*., 2009; Jimbo and Jawad, 2018). This test is used to examine the null hypothesis of constant variance. A p-value of 0.05 or less indicates that the issue of heteroscedasticity exist. Thus, the model is considered homoscedastic and we fail to reject the null hypothesis of constant variance.

Multicollinearity

Multicollinearity problem exists when two independent variables are highly correlated. To check for the existence of multicollinearity problem in the model, drawing the correlation matrix of all independent variables is the simplest approach. However, researchers use VIF as a more formal test for detecting multicollinearity (Claver *et al*., 2018). The Variance Inflation Factor (VIF) test gives a more accurate information about multicollinearity by providing an index that gages how much the variance of coefficient in the regression model is increased because of collinearity (Kutner *et al*., 2005). If the VIF value reaches 10, the multicollinearity is considered problematic. In that case, one of the two highly correlated variables is omitted from the model.

Autocorrelation

The other assumption of an OLS regression is independence of residuals. The residuals should be independent. Autocorrelation problem exists when the independent variables are time-dependent. To evaluate the model for autocorrelation, the easiest approach is to plot the residuals against time in a scatter plot.

Pearson Correlation Coefficient

It is used to show the direction and strength of linear correlation between two continuous variables. It has a value between -1 and +1, where 1 indicates perfect positive linear correlation, 0 indicates no linear correlation and -1 indicates perfect negative linear correlation. As in this study the data to be analyzed are samples, the following formula is used to calculate the Pearson correlation (Kutner *et al*., 2005):

$$r_{xy} = \frac{n\sum_1^n x_i y_i - \sum_1^n x_i \sum_1^n y_j}{\sqrt{n\sum_1^n x_1^2 - n\left(\sum_1^n x_1\right)^2}\sqrt{n\sum_1^n y_i^2 - \left(\sum_1^n y_i\right)^2}} \tag{5}$$

where, r_{xy} denotes the pearson correlation coefficient, n is number of observations, x_i and y_i represent the i^{-th} observations in x and y datasets.

Analysis and Result

In this section, the data are analyzed, tested and validated. These steps are important and ultimately used for hypothesis testing.

The data analysis is preceded with an overview of the survey respondents. The hypotheses are tested by employing regression and correlation methods. The OLS Regression is used when the dependent variable is continuous and Pearson correlation methods is used to find the direction and strength of pairs of variables. For the purpose of statistical data analysis, we use STATA software to analyze the collected data.

Overview of Collected Data

Response Rate

The response rate was 50% (20/40). That is to say, of the 40 sets of questionnaires distributed to the public and private organizations in Kabul, they returned 20 duly filled sets. About 50% of questionnaires were either returned defected or not returned at all (Table 1).

Profile of Respondents

The respondents were mainly selected among IT experts from governmental and non-governmental organizations. The detailed information or profile of the respondents were also collected. The demographic information gender, age, employment status and education level of the respondents were also added.

The simplified information on the collected data is presented in the Table 2.

Reliability Analysis

To test the reliability of the variables, *Cronbach's Alpha* was calculated for technical questions of both questionnaires. The Cronbach's Alpha coefficient is higher than 0.69, which means that the data are reliable. The Table 3 shows the result of Cronbach's Alpha test.

The R_2 is a statistical measure that o ho closed the data complexity and its goodness of. Table 5 shows that approximately 50% the observed variation can be explained by the model s inputs.

Regression Analysis

To have better and scientifically acceptable result from the regression models, the following pre- and post-regression tests are conducted for each model.

Before Running the Regression Model Winsorization

Outliers are detected and winsorized to reduce their effects on outcomes of the regression models. Akaike Information Criteria (AIC): It is conducted to include independent and control variables in the model by balancing between model complexity and its goodness of fit (balancing between over-fitting and under-fitting)

Table 1: A survey distributed to more than 40 organization in Afghanistan, which includes private and government sectors. The following table shows the result of survey and the percentages of correspondence

Description	Number	Percentage
Questionnaires distributed	40	100
Questionnaires returned	20	50
Questionnaires unreturned	20	50

Table 2: Respondent's Profile

Description	Number	Percentage
Priv. organization	12	25.6
Publ. organization	8	24.4
Missing	20	50
Total	40	100

Table 3: Results of Cronbach's alpha test (refer to reliability analysis subsection)

Questionnaire	Number of items	Cronbach alpha	Reliability
Organization	40	0.76	Acceptable

Table 4: Data lost significally depends on direct responsibility, ISP audit and ISP standard (refer to regression analysis and results subsections)

Data lost	Coef.	Std. Err.	t	P-value
Direct responsibility	-4.46984	2.076998	-2.15	*0.0.31
ISP_Audit	-5.124083	1.991061	-2.57	*0.010
ISP_phys_security	0.6394064	1.562187	0.41	**0.682
ISP_standard	2.509436	1.211245	2.07	*0.038
Business type	-2.066337	1.386575	-1.49	**0.136
ISP_encreption	-1.895273	1.2891	-1.47	**0.141
Security Breach	-1.18446	0.8563653	-1.38	**0.167

Note. *p<0.05, **p>0.05

Table 5: Model summary test (refer to reliability analysis subsection)

Observations	R_2	Adjusted R_2	F	Prob>F	df
263	0.49	0.46	9, 153	0.001	29

After Running the Regression Model

The OLS underlying assumptions (normality, linearity, multicollinearity, heteroscedasticity, autocorrelation) are checked. If the assumptions are not fulfilled, data are transformed and smoothed.

Linear Regression Analysis

OLS multiple regression model is conducted to analyze the relationship between the dependent variable and a set of predictor variables. The variables were transformed by taking their square roots because *homoscedasticity* and *normality* assumptions were not fulfilled. The results of testing the assumptions of OLS regression after transformation process follow the regression model presented below.

Dependent Variable

Data Lost y

This variable shows the loss of information related to security and privacy parameters. It has to be presented or designed before filling the questionnaire. It is used as a dependent variable to represent the effect of the predictor variable on security and privacy in cloud computing.

Predictor Variables

x_1: *Direct responsibility*
x_2: *ISP_Audit*
x_3: *ISP_Phys_Security*
x_4: *ISP_Standard*
x_5: *Business Type*
x_6: *ISP_Encreption*
x_7: *Security Breach*

Results

This research investigates the factors that affect the data loss in cloud computing. The ultimate goal in this study is to detect which of the parameters related to security and privacy might affect the cloud computing as a whole. We found using statistical techniques, that, the direct responsibility, isp_audit and phys isp standard were the most significant variables affecting data loss in cloud computing. Other variables such as ISP_Standard, business type, isp_encreption and security breach did not affect the data loss at all. In this research work we have proposed six models and used the AIC to select the best model among them. The best model here corresponding to the model with the smallest AIC value.

Discussion and Recommendation

In this section, the findings of study, which were presented in the previous section, are explained and discussed in details. Some comparative analysis is made to compare the findings of this study with findings of other similar studies. In this section we will present the discussion of our results, some limitations and provide some recommendations for future research directions and developments.

Discussion

The study of the impact of security and privacy in cloud computing has allowed us to assess the impact of several factors that have been developed in this study. When exploring the literature review, one of the main limitations we faced in this research was the absence of relevant literature available.

Also we discovered that such novel direction in the research in cloud computing has not been carried out in Afghanistan. Two organizations were explored in this study, namely public and private organizations. We observed that private organizations had better response rates compare to the public ones. They also have better acceptance and investments in privacy and security issues in order to optimize both their management and profit margins.

Recommendation to Governments and Private Organizations

At programming and policy levels, based on the findings of the study, there are a number of actions recommended that the public and private organizations and industry owners can take into consideration to make use of the advantages of security and privacy issues in cloud computing and in the meantime prevent or/and reduce the negative effects of the data loss on their management systems. These actions are summarized in three points as follows.

The Public Organizations

In case they do not have sensitive information should use the cloud computing services in order to have stable services, availability for better services for clients, high security protection and consumer satisfactions.

The Private Organizations

By using the cloud services are able to have 99.9% accessibility to the data in the cloud with 0 downtime in their businesses. All private sectors need to have accurate and accessible information, a few second

downtime can affect their business and generate loss of massive amount of money and services. Besides, these organizations do not need to invest lots of their money in order to protect their security and make adequate planning for the disaster recovery. Cloud computing provides the services that most organizations need with high security and low price that, the organizations by themselves will not be able to have.

Conclusion

Cloud computing is a model that helps to speed up and increase the flexibility of data management with reduced cost. It is undeniable that cloud computing has brought us lots of benefits and becoming more popular nowadays. Many large companies start using cloud service in their business. While the cloud computing is widely used, the security and privacy become a concern to everyone who uses cloud services. There are lots of security issues arising continuously while there are also improvements as well on the security models of the cloud services provided. Despite the increasing use of the cloud services, the user should use the cloud services provided wisely in a way that always ensures good security practices so that these technologies have the potential to bring the information technology to the next level. Cloud computing might help us to separate the software from the hardware as more as technologies are used as service using cloud and software might have a highly abstract space with the computer hardware. It is expected that this study provides some foundations in regards to issues and challenges of security and privacy in cloud computing.

Whatever we think about today, the security and privacy of cloud computing challenges have become critical for us. But what we can do is just to follow up the service provider and their guidance on the background they are working on everything regarding security and privacy. We have to select things that are important to us and they will be included in Service Level Agreements (SLA) that is contracted between consumer and provider. Whatever is mentioned, the cloud provider will be responsible for it. The consumer can claim to the cloud service provider according to a contract he already had with them.

Acknowledgement

This study was stimulated by the series of discussions in the weekly research discussion meetings our laboratory and a great conversation with few colleagues at NACA 2016, Niigata, Japan. The research presented in this study is supported by the Sakura Research and Waseda University, Tokyo, Japan. We thank all colleagues for fruitful discussions. We equally thank Hamidullah Hamidy for reviewing and editing the final version of this manuscript. This publication reflects only the authors' views and any remaining mistakes are ours.

Author's Contributions

Jimbo Henri Claver: Designed, presented and wrote the model used in this study. As principal investigator, he supervised the project as a whole and drafted the manuscript. All authors read and approved the final draft of the manuscript.

Edris Hamraz: Designed the questionnaires and did all the data analysis work.

Takeru Suzuki: Contributed with his expertise in this research work and provided interesting ideas for improving the quality of the manuscript.

Ngongo Seraphin: Contributed in designing the questionnaires and developing various statistical models we used for testing and validation.

Andjiga Gabriel: Assured the high quality of this research work as a whole and guided all of us with his outstanding experience in applied mathematics.

Etoua Magloire: Contributed in reviewing the article for significant intellectual content and scientific quality with his exceptional expertise in predictive modelling.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

Availability

The data collected and used in the research is available upon request.

References

Bin, W., Yuan, H. H., Xi, L. X., & Min, X. J. (2009, October). Open identity management framework for SaaS ecosystem. In 2009 IEEE international Conference on e-Business Engineering (pp. 512-517). IEEE.

Bozdogan, H. (2000). Akaike's information criterion and recent developments in information complexity. Journal of mathematical psychology, 44(1), 62-91.

Claver, J., Hamraz, E., Azimi, J., & Owona, C. (2018). Optimal Allocation of QoS and Web Services in Cloud Computing. American Journal of Information Systems, 6(1), 23-28.

CSA, 2010. Cloud security alliance. http://www.cloudsecurityalliance.org/

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS quarterly, 1165-1188.

Elizabeth, F., & Vadim, O. (2007). Web application scanners: Definitions and functions. HICSS 2007, 280b-280b.

ENISA, 2010. Cloud computing: Benefits, risks and recommendations for information security. http://www.enisa.europa.eu/act/rm/files/deliverables/cloud->computing-risk-assessment

Grobauer, B., Walloschek, T., & Stocker, E. (2010). Understanding cloud computing vulnerabilities. IEEE Security & privacy, 9(2), 50-57.

Holstein, D. K., & Stouffer, K. (2010, January). Trust but verify critical infrastructure cyber security solutions. In 2010 43rd Hawaii International Conference on System Sciences (pp. 1-8). IEEE.

Jensen, M., Schwenk, J., Gruschka, N., & Iacono, L. L. (2009, September). On technical security issues in cloud computing. In 2009 IEEE International Conference on Cloud Computing (pp. 109-116). IEEE.

Jimbo, H. C. (2019). Introduction to Security and Privacy in Cloud Computing. 1st Edn., Lambert Academic. pp: 80.

Kandukuri, B. R. R. Paturi, V., & Rakshit. A. (2009). Cloud Security Issues. In 2009 IEEE International Conference on Services Computing (Vol. 10).

Jimbo, H. C., & Jawad, A. (2018). Multi-factor and Dimensional Approach in Data Analysis. 1st Edn., Lambert Academic.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied linear statistical models (Vol. 5). New York: McGraw-Hill Irwin.

Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

Microsoft. (2010). Website: Multi-tenant data architecture. http://msdn.microsoft.com/en-us/library/aa479086.aspx

NIST. (2010). National Vulnerability Database (NVD). http://nvd.nist.gov/home.cfm

OWASP. (2010). The ten most critical web application security vulnerabilities. http://www.owasp.org/index.php/OWASP_Top_Ten_Project

Peter, M., & Tim, G. (2009). The NIST definition of cloud computing. http://www.wheresmyserver.co.nz/storage/media/faq-files/cloud-def v15

Popović, K., & Hocenski, Ž. (2010, May). Cloud computing security issues and challenges. In The 33rd international convention mipro (pp. 344-349). IEEE.

Ristenpart, T., Tromer, E., Shacham, H., & Savage, S. (2009, November). Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In Proceedings of the 16th ACM conference on Computer and communications security (pp. 199-212).

Subashini, S., & Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. Journal of network and computer applications, 34(1), 1-11.

Zhang, W. (2010, April). Integrated security framework for secure web services. In 2010 Third International Symposium on Intelligent Information Technology and Security Informatics (pp. 178-183). IEEE.

Tables for Models

We present the output of each model respectively in the Tables 1 to 6.

Table 1: Model one

. glm Data _lost Direct responsibility ISP_Audit ISP_ Phys_Security business type internet Attack				
Iteration 0: Log likelihood=-26395036				
Generalized linear models				
Optimization:	ML	No. Of obs	=	17l
		Residual df	=	11
		Scale parameter	=	2.0192777
Deviance	= 22.21204275	(1/df) Deviance	=	2.019277
Pearson	= 22.21204275	(1/df) Pearson	=	2.19277
Variance function:	V(u) = 1	(Gaussian)		
Link Function	: g(u) = u	(Identity)		
Log likehood	=-26.39503581	AIC	=	3.811181
		BIC	=	-8.953304

Table 2: Model two

.glm Data _lost Direct responsibility ISP_Audit ISP_ Phys_Security business type internet Attack				
Iteration 0: Log likelihood=-26.029701				
Generalized linear models		No. Of obs	=	17
Optimization:	ML	Residual df	=	11
		Scale parameter	=	1.934326
Deviance	= 21.27758257	(1/df) Deviance	=	1.934326
Pearson	= 21.27758257	(1/df) Pearson	=	1.934326
Variance function: V(u)	= 1	(Gaussian)		
Link Function	: g(u) = u	(Identity)		
Log likelihood	=-26.02970113	AIC	=	3.7682
		BIC	=	-9.887764

Table 3: Model three

.glm Data _lost Direct responsibility ISP_Audit ISP_ Phys_Security business type internet Attack				
Iteration 0: Log likelihood = -26.929169				
Generalized linear models				
Optimization:	ML	No. Of obs	=	17
		Residual df	=	12
		Scale parameter	=	1.971051
Deviance	= 23.645261776	(1/df) Deviance	=	1.971051
Pearson	= 23.65261776	(1/df) Pearson	=	1.971051
Variance function: V(u)	= 1	(Gaussian)		
Link Function	: g(u) = u	(Identity)		
Log likelihood	= - 26.92916891	AIC	=	3.756373
		BIC	=	-10.345994

Table 4: Model four

. glm Data _lost Direct responsibility ISP_Audit ISP_ Phys_Security ISP _Standard business type ISP_Encreption.				
Iteration 0: Log likelihood = -24.429006				
Generalized linear models				
Optimization:	ML	No. Of obs	=	17
		Residual df	=	10
		Scale parameter	=	1.762533
Deviance	= 17.62532933	(1/df) Deviance	=	1.762533
Pearson	= 17.62532933	(1/df) Pearson	=	1.762533
Variance function: V(u)	= 1	(Gaussian)		
Link Function	: g(u) = u	(Identity)		
Log likelihood	= - 24.42900643	AIC	=	3.69753
	BIC	=	-10.7068	

Table 5: Model five

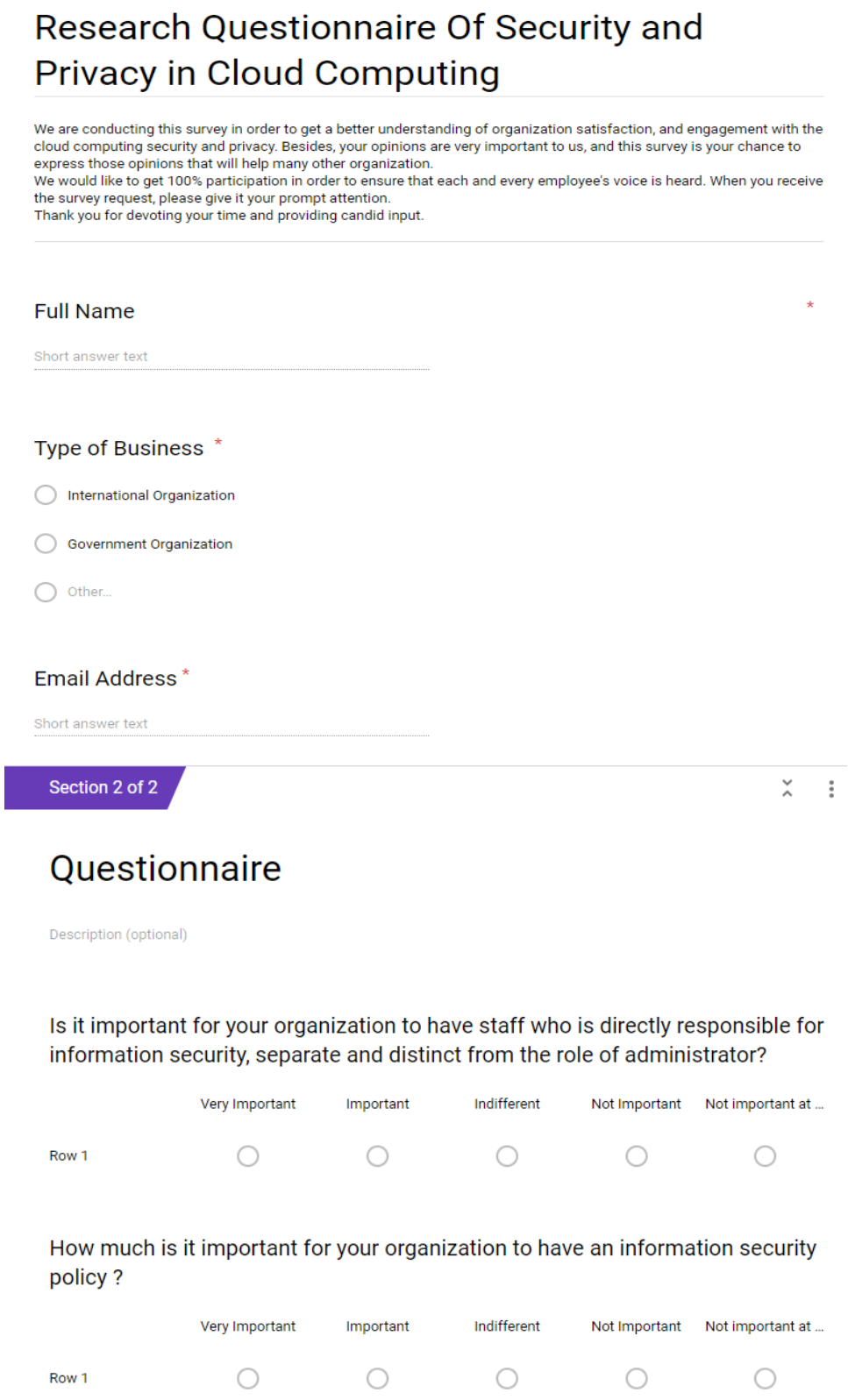
			Number of obs	=	17	
			LR chi2(7)	=	20.36	
			Prob > chi2	=	0.0048	
Ordered logistic regression			Pseudo R2	=	0.4139	
Log likelihood = -14.417809					(95% Conf. Interval)	
Data Lost	Coef.	Std. Err.	Z	P>|z|	----------	----------
Direct Responsibility	-4.467948	2.076998	-2.15	0.031	-8.538789	-3.3971074
ISP_Audit	-5.124083	1.991061	-2.57	0.010	-9.02649	-1.221676
ISP_Phys_Security	.6394064	1.562187	0.41	0.682	-2422424	3.701237
ISP_Standard	2.509436	1.211245	2.07	0.038	0.1354402	4.883432
BusinessType	-2.066337	1.386575	-1.49	0.136	-4.783973	0.6512999
ISP_Encreption	-1.895273	1.2891	-1.47	0.141	-4.421863	0.6313158
Security Breach	-1.18446	0.856363	-1.38	0.167	-2.862905	0.4939847

Table 6: Model six

.glm Data _lost Direct responsibility ISP_Audit ISP_ Phys_Security ISP_Standard business type internet Attack				
Iteration 0: Log likelihood = -22.793646				
Generalized linear models				
Optimization:	ML	No. Of obs	=	17
		Residual df	=	9
		Scale parameter	=	1.615617
Deviance	= 14.54055512	(1/df) Deviance	=	1.615617
Pearson	= 14.54055512	(1/df) Pearson	=	1.615617
Variance function: V(u)	= 1	(Gaussian)		
Link Function	: g(u) = u	(Identity)		
Log likelihood	= -22.79364567	AIC	=	3.622782
	BIC	=	-10.95836	

DOI: 10.3844/jcssp.2020.1625.1638

Structure of Questionnaires: The design of questionnaires requires the information below.



Research Questionnaire Of Security and Privacy in Cloud Computing

We are conducting this survey in order to get a better understanding of organization satisfaction, and engagement with the cloud computing security and privacy. Besides, your opinions are very important to us, and this survey is your chance to express those opinions that will help many other organization.
We would like to get 100% participation in order to ensure that each and every employee's voice is heard. When you receive the survey request, please give it your prompt attention.
Thank you for devoting your time and providing candid input.

Full Name *

Short answer text

Type of Business *

International Organization

Government Organization

Other...

Email Address *

Short answer text

Section 2 of 2

Questionnaire

Description (optional)

Is it important for your organization to have staff who is directly responsible for information security, separate and distinct from the role of administrator?

	Very Important	Important	Indifferent	Not Important	Not important at ...
Row 1					

How much is it important for your organization to have an information security policy ?

	Very Important	Important	Indifferent	Not Important	Not important at ...
Row 1					

Does for your organization, it is important to follow any recognized standard for information security policy

- [ ] Web Trust
- [ ] ISO/IES 27001
- [ ] NIST
- [ ] Other...

If you have performed a third-party audit of any aspect of your data security, is your organization willing to supply a written copy of important audit report for a review?

	Very Important	Important	Indifferent	Not Important	Not important at ...
Row 1	○	○	○	○	○

Does your organization's information security policy include specific provisions for responding to security breaches and is it important for them?

	Very Important	Important	Indifferent	Not Important	Not important at ...
Row 1	○	○	○	○	○

Does physical security is important for information security policy of your organization?

	Very Important	Important	Indifferent	Not Important	Not important at ...
Row 1	○	○	○	○	○

Do your organization has been experiencing any internet attack in the last year

	Yes in 2017	Yes, 2016, 2017	No experience	Don't know
Row 1	○	○	○	○

Have you been experienced losing data in your organization after an attack?

- [ ] 1- 20 %
- [ ] 20 - 40 %
- [ ] 40 - 60
- [ ] 60 - 80 %
- [ ] 80 - 100
- [ ] No Data Loss
- [ ] Other...

DOI: 10.3844/jcssp.2020.1625.1638

Appendix C. Response Rates of the International and Government Organizations in Afghanistan