Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.11, No.5, 2020
www.iiste.org

Survey of Data Confidentiality and Privacy in the Cloud Computing Environment

Michael Alberto Larbitey Lamptey[1*] Guangping Zhou[2]
School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China

Abstract
The objective of this research is to develop a scheme for improving cloud data confidentiality. A considerable number of people are sharing data through third-party applications in the cloud computing environment. According to reviewed literature, it has been realized that data security and privacy were the key challenges to the wider adoption of cloud services with insider threats being the most prevalent. Similarly, our online survey indicated that 53.3% of the respondents citing insider breaches as the main threat to their organizational data. The survey also confirmed that data security and privacy is one of the greatest barriers to the adoption of cloud services in their organization. Noting the flaws of Attribute-Based Encryption (ABE) and Identity-based encryption (IBE), and with the growth of computing power, applications are constantly being developed which makes them vulnerable to attacks. Since data confidentiality is essential in the provision of information security in the cloud, this paper suggested the development and the deployment of a hybrid attribute-based re-encryption scheme, which is a scheme that bridges the ABE and IBE, to secure data in the cloud computing environment.
Keywords: Encryption, Cloud Computing, Data, confidentiality, Privacy
DOI: 10.7176/CEIS/11-5-03
Publication date: September 30th 2020

1. Introduction

Widespread confusion has become lucid recently concerning the meaning of cloud computing (Daylami, 2016). According to Kepes (2017), cloud computing is amorphous since it does not settle on one thing. Generally, cloud computing has three main forms which are Infrastructure as a service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) (Tai & Bermbach, 2017). Data confidentiality and privacy in the cloud computing environment is an emerging notion whose demands continue to shift as they address ancient vulnerabilities. Cloud data privacy as an element of security is constantly evolving owing to an ever-changing internet setting.

In this information age, one of the greatest problems facing organizations is the failure to ensure data confidentiality and privacy and many organizations are not willing to adopt the unprecedented online demand services spearheaded by cloud computing. Evidence is demonstrating that despite cloud computing being seen as a major business avenue for the next years, migration to the cloud paradigm is hampered by concerns on security (Coppolino, Antonio, Mazzeo, & Romano, 2017). Even though cloud computing provides less cost and less resource management, it has some security threats (Rao, 2016). Referencing to the 2017 private data leak statistics, they are sufficient to demonstrate that a single event can have a huge effect on the reputation and finances of a firm. The Equifax leak is the most recent example that affected 146.6 million people, having names, birthdates, Social Security numbers, and addresses stolen. But this scenario is just a one drop amidst a sea of poorly secured private information cases in the cloud computing environment. Coming from a public database exposing 198 million American voter records, and moving to FedEx leaking which exposed 119 000 scanned passports, driver licenses, and other personal documents on a publicly accessible server, the list seems to be endless.

These companies had no intention of misusing private information even if they were guilty of not implementing data confidentiality and privacy strategies. The difficult reality is that, given the evolution of cloud computing, cybercrime, and the evolving data protection regulations around the globe, it is quite apparent that each business should develop (and follow) clear guidelines for data confidentiality and privacy. However, this will involve a major shift in corporate culture for most organizations, which is not possible without a corporate-specific approach that recognizes the need for data privacy in the cloud computing environment.

To protect this valuable data, several pairing-based cryptosystems have been developed. These include the development of prominent cryptographic techniques to provide privacy and fine-grained access control in cloud computing such as Attribute-Based Encryption (P, P, & Alphonse, 2018). However, these schemes are not flawless which may result in the aforementioned data leaks. ABE requires that a priori access policies should be determined for various ciphertexts upon encryption but these policies will become obsolete after some time (Deng, Qin, Wu, Guan, & Zhou, 2020). Another condition is that ABE decryption complexity increases linearly in the number of attributes and eventually incurs considerable amounts of computations that are unfavorable for resource-limited users (Deng et al., 2020). Thus, these conditions render the ABE scheme inappropriate and insufficient for some sophisticated applications.

2.0 Scenario for Study

Due to the rapid advancement of the online lifestyle in this information age, cases of hacking into individual email accounts, and organization servers are increasing. The threat landscape is changing with occurrences of unauthorized access becoming more dominant, with 60% of all attackers being insiders (Feldman, 2012). Cloud computing is rife with a history of leaks, both accidental and deliberate that has led to the recognition of the privacy risks of cloud deployment. Due to its unprecedented and overwhelming benefits, cloud computing has been considered beyond reasonable doubt as one of the most influential technologies in the information era. The global configuration of Post Fordism has facilitated the free movement of capital, information, and ideas across boundaries through the world wide web and cloud computing. This implies that capitalists and cloud computing users lose the direct control of their data and completely rely on cloud technologies to manage their data. Moreover, the present cloud storage confidentiality techniques are still evolving and some suffering from still-birth. As they evolve so are the threats too and this has posed security and privacy challenges. With this regard, significant security and privacy concerns are being raised and many organizations are not willing to adopt the cloud technologies. Therefore, it is desirable to develop an outlook for improving data confidentiality and privacy in the cloud computing environment

3.0 Literature Review

3.1 Introduction

According to Britt (2019), cloud safety relates to protecting the running apps, stored information, and cloud transaction processing. “The proponent further alluded that the additional growth attacks are in exponential factors. This, therefore, requires ongoing knowledge of how and when to secure stored information and apps in the cloud environment. According to NIST (2004), continuous monitoring involves an ongoing awareness of stored information security, threats, and vulnerabilities that will facilitate risk-based decision making. It has been noted by Britt (2019) that there has been an accumulation of malware attacks in the past 2 years as compared to the past 18 years. While operations in the cloud environment have been met with great expectations in terms of scalability, agility, and efficiency, security concerns evolving around cloud computing have not been fully addressed. Presently, no framework can allow CSUs to evaluate CSPs based on their ability to meet the customer’s security requirements (Rizvi, Ryoo, Kissell, Aiken, & Liu, 2017). What then it simply implies is that there is an outcry worldwide on the need to secure endpoint devices and valuable information needed by various online users. Thus, the data confidentiality and privacy framework in the cloud computing environment need to address the security risks and concerns if end users are to benefit from the potential gains of low-cost computing. The following security concerns are to be observed in the development of the data confidentiality and privacy outlook in the cloud computing environment.”

3.1.1 Inside Attacks

Cloud computing is a multitenant-based model under the single management domain of the provider (Gupta, Laxmi, and Sharma, 2014). “Other proponents such as Seema and Shaikh (2014) asserts that cloud computing represents the relocation of computing power from the organizational premises to several computers or nodes transmitting data in the cloud. While the former proponents (Gupta, Laxmi, and Sharma) stresses cloud computing based on having a single managed domain, the latter contravened the notion by indicating the shift of ownership to renting and this has raised concerns especially for organizations that have no data confidentiality and privacy framework to regulate the use of data in the cloud computing environment. Moreover, there has been some collaboration between subscribers and cloud service providers in data management within the cloud computing environment but the reality still stands that malicious insider attacks are not even detected in some organizations.”

3.1.2 Outsider Attacks

Seema and Shaikh (2014) refer to a cloud as a virtualized server pool that provisions considerable computing power to end-users. “Contrarily, multiple organizations concentrated on a Cloud Service Provider may have more appealing benefits than a single organization (Crowe, 2012). While the authors in this regard elaborate about the power of virtual integration provided by the Cloud Service Providers, this integration has created the best target for attacks due to the publicity created. Thus, the argument raised questions on the need for data confidentiality and privacy scheme. The cloud computing environment encompasses additional programming interfaces as compared to the private networks and this makes them vulnerable to hackers and attackers who can then obtain the advantage of exploiting the API flaw (Gupta, Laxmi, and Sharma, 2014). Generally, the conception makes cloud data susceptible to risk and attacks and hence the need for a scheme.

3.1.3 Data breaches

According to the Cloud Security Alliance (2016), a data breach is an occurrence in which information that is vital and private is observed, released, used by an unauthorized person, or stolen. There are a handful of data breaches as serious challenges in the cloud computing environment (Samson, 2013). Data breaches and among other methods such as SQL injection and virtual machine side-channels have been used by malicious attackers to get the decryption and encryption keys that are used by other machines on the same server. Data breaches that have been

the traditional way to attack machines with a networked environment in the past have been brought to the cloud computing environment. Ideally, this clearly shows that some challenges prevalent in the traditional computing environment will certainly find their way to the new cloud computing environment. Thus, this saturates the already vulnerable environment with fresh threats. Moreover, as Samson has indicated that the remedies put in place may exacerbate the others, this will, therefore, call for trade-offs between information sensitivity and security to be stored in the cloud environment.

3.1.4 SQL Injection

The Standard Query Language (SQL) and XML are the two mainly used database types. These types of databases are highly useful in storing data from their associated websites. SQL is useful in querying relational databases and it uses command and control language while the XML is specifically for querying XML databases. In SQL injection, the attack sends mischievous commands direct to the database by sneaking through illegal means. An SQL injection is a deadly attack that can be employed by criminals as well as the state to achieve their unethical goals. This form of attack is becoming a popular choice for hackers because of its nature to facilitate the execution of data or its interpretation in an unpredicted manner. There are several forms in which SQL injection can take.

These include head injection, log injection, full path disclosure, and cross-site scripting (XSS). Because of the numerous types they can take, injections are the most prominent attacks proliferating on the internet. The question that comes into play is how the various forms of SQL injection can be prevented.

3.2 Protection of Data Privacy through Encryption

The traditional encryption schemes such as symmetric and asymmetric schemes are not suitable to provide access control due to lack of flexibility and fine-grained access control. One of the prominent cryptographic techniques to provide privacy and fine-grained access control in cloud computing is Attribute-Based Encryption.

3.2.1 Attribute-Based Encryption

Attribute-based encryption is one of the most prominent technique to provide privacy and fine-grained access control as compared to the traditional cryptographic schemes such as asymmetric and symmetric schemes. The traditional schemes do not give full access control and data privacy due to a lack of fine-grained access control and lack of flexibility. ABE was first proposed by Sahai and Waters (2005) and the proponents suggested that a public key of the one-to-many algorithm protects the data in the cloud. Data encryption as suggested in this scenario is based on the set of attributes where three entities are considered. These entities are the Data User, Authority, and the Data Owner. The certified Authority initially generates a public key which it sends to the data owner for encryption and also generates the master secret key. Ideally, the Authority produces the user's secret key with the master key according to the attributes. The associated Data owners receive the public which they then use to encrypt data using attributes and store them in the cloud. The algorithm is as follows:

Setup:

Let ℓ: G1 X G1 G2 denote the bilinear map.

The certifying authority generates the public parameter Public Key (PK) and Master Secret Key (MSK) using the random value X_1, X_2, X_N, and Y from Zn

- *Key Generation:*

The Authority generates the user's secret key (S) associated with every user. Given that the q (0) = y, the user's secret key S is denoted by:

where J AU; AU UA

- *Encryption:*

The encryption algorithm produces a *ciphertext*. The message M in this regard is encrypted based on the PK and set of attributes that are used as identified by the data owner. The encrypted text E is defined as follows:

- *Decryption:*

The data user decrypts the message from the ciphertext using the secret key S, generated with the identity AU. The ciphertext is generated with the identity () and the decryption is only possible if ().

The message M is computed as follows:

M ¼ E0=Yc

3.2.2 Identity Based Encryption

IBE is another prominent scheme that provides fine-grained access control and is also characterized by four randomized algorithms which are Setup, Extract, Encrypt, and Decrypt.

- *Setup*:

The setup takes a security parameter k and returns params (system parameters) and master-key. The system parameters include a description of a message space *M* and a description of a Ciphertext space C. Intuitively, the system parameters will be publicly known, while the master-key will be known only to the "Private Key Generator" (PKG).

- *Extract:*

The extract takes input params, master-key, and an arbitrary ID 2 f0; 1g_, and returns a private key d. Here ID is

an arbitrary string that will be used as a public key, and d is the corresponding private decryption key. The Extract algorithm extracts a private key from the given public key.

- *Encrypt:*

The encrypt algorithm takes as input params, ID, and M 2M. It returns a ciphertext C 2 C.

- *Decrypt*:

The algorithm takes as input params, C 2 C, and a private key d. It returns M 2M. These algorithms must satisfy the standard consistency constraint, namely when d is the private key generated by algorithm Extract when it is given ID as the public key, then 8M 2M: Decrypt (params; C; d) = M where C = Encrypt (params; ID; M)

4.0 The Methodology

4.1 Research Design

The study used a combination of both descriptive research and explanatory research techniques. Descriptive research tries to elucidate the current gaps in the cloud computing environment and how these gaps will be addressed. Likewise the explanatory looked on why the world is in pursuit of remedies in data confidentiality and privacy. Thus this was purposefully constructed to examine the trends and how they informed the researcher in addressing the questions inherent. A survey was carried out using the Open Data Kit platform with systems administrators of companies that deal with cloud services. The survey was aimed at obtaining general information about cloud computing within their respective organizations before indulging in a series of experiments to evaluate the performance of the hybrid attribute-based re-encryption scheme schemes in practice. Random sampling was used because each elementary object in the targeted population had an equal chance of being selected. The sampled population consisted of employees from various mobile telecommunications companies, banking services, health services, and supermarkets that use cloud computing services. In this scenario where random sampling was used, Slovin's formula was appropriate. Ellen (2016) noted that the formula is used when nothing regarding the behavior of a population is known at all. The formula is represented as shown below:

n=N/ (1 + Ne2)

Where:

n=number of sample size
N= total population
e=confidence level

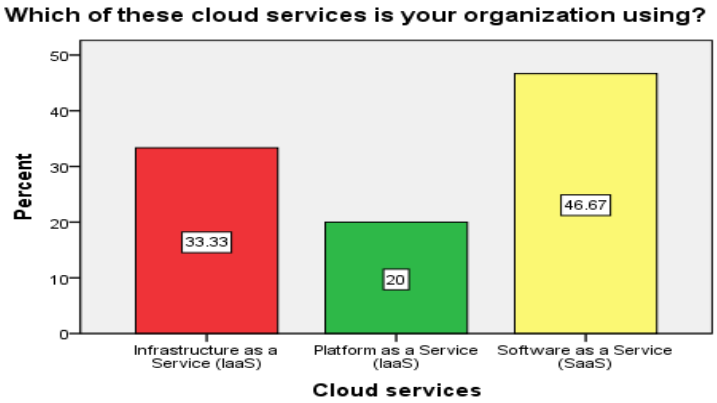
The formula above is very useful since it does not involve the use of the mean in the determination of the sample size as required by other formulas. With this regard, confidence levels of e =0.02 were used accordingly and accuracy of about 98% was achieved that reflected the true picture of the threats to data confidentiality and privacy within the cloud computing environment.

5.0 Findings and Analysis

The researcher surveyed various administrators who deal with various cloud services. The aim was to gather relevant statistical data and general information concerning data confidentiality in the cloud computing environment before investigating the greatest challenges or barriers in the adoption of cloud computing services. Moreover, the research study would also endeavor to identify the greatest threat to data confidentiality and privacy in the cloud computing environment before developing a data privacy scheme. A total of 30 respondents participated in the survey which was administered.

5.1 Cloud Services Adopted by Organizations

Kepes (2017) noted that cloud computing is amorphous since it does not settle on one thing. Generally, cloud computing has three main forms which are Infrastructure as a service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). To determine the cloud services adopted by organizations, respondents were asked based on these three categories. .The figure below shows more use of Infrastructure as a service (IaaS**).**



Source: Research data 2019

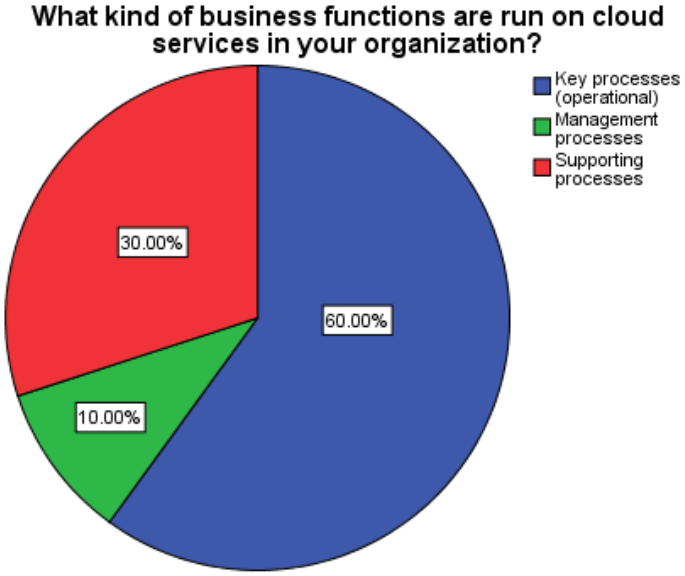
Figure 5.1. More use of Software as a service (SaaS)

The results from the respondents indicated that most organizations adopted the use of cloud services mainly Software as a Service (SaaS) and Infrastructure as a Service (IaaS). Ideally, as the systems become more complex, the traditional cryptosystems become as symmetrical encryption does not provide fine-grained privacy and access control.

A more sophisticated paradigm to encourage the adoption and use of cloud services is to reduce the vulnerabilities by implementing new cryptosystems that enhance data privacy and confidentiality. Technically this was suggested to be in the form of a flawless hybrid attribute-based encryption that combines IBE and ABE.

5.2 Business Functions Supported by the Business

Business functions are supported in the cloud and have a significant impact and strategic advantage offered to the organizations.



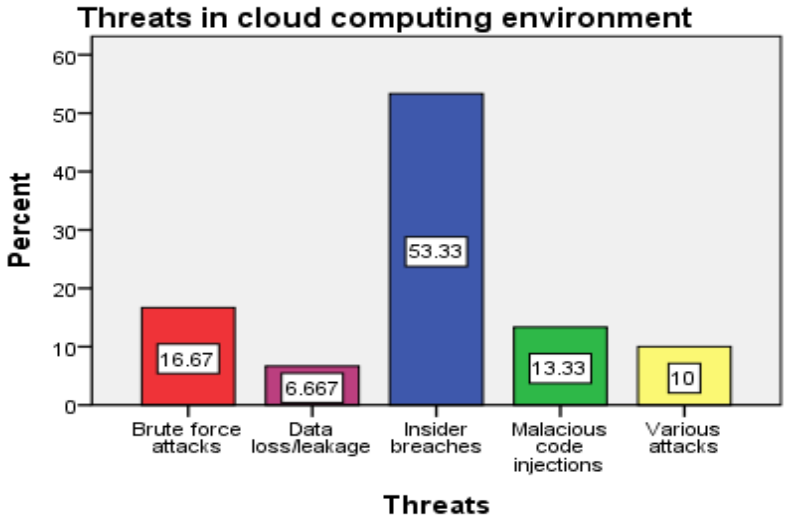
Source: Research data 2019

Figure 5.2 Business Functions Supported

Figure 5:2 above illustrates how cloud computing services support key processes (operational) and supporting processes (like human resource management) in day to day organizational operations. As noted, most organizations use cloud computing applications to support key or operational processes. Most of these applications have implemented traditional IBE schemes such as those suggested by Bone and Franklin (2001). Whenever a user or a key process is revoked from the system, there must be a revocation mechanism that caters to such a change to prevent an attack. Thus this has not been possible with these schemes resulting in data loss such as the Equifax leak.

5.3 Cloud Computing Threats

Threats in the cloud computing environment are inherent due to poor implementation and have been exacerbated by the non-scalability of these traditional schemes.



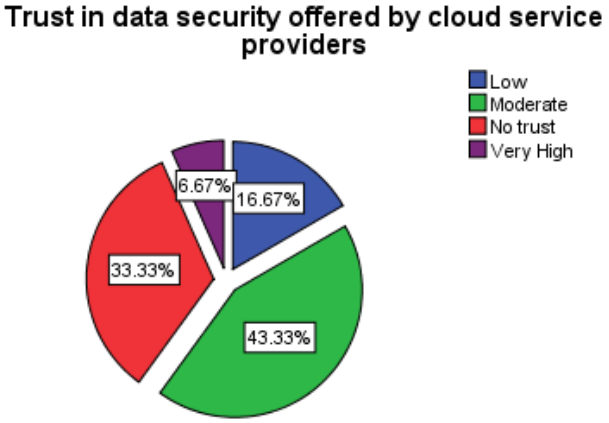
Source: Research data 2019

Figure 5.3 Threats in the Cloud Computing Environment

The results show that insider breaches were the main forms of vulnerabilities to organization data within the cloud environment. As shown above in figure 5.3. This implies that the Data breaches which have been the traditional way to attack networked machines have been brought to the cloud computing environment. This can be through, co-residence attacks where a user's information in one VM can be accessed (stolen) or corrupted through side channels by a malicious attacker's VM co-residing on the same server (Levitin, Xing, & Dai, 2018). The malicious attackers obtain the decryption and encryption keys that are generated by the traditional cryptosystem algorithms.

5.4 Trust in Cloud Computing Services

The survey also confirmed that despite the rapid growth of cloud computing, data confidentiality and privacy in the cloud servers were still a big challenge and there is a need for continued research on how best to secure data within that environment.



Source: Research data 2019

Figure 5.4 Trust in Cloud Computing Services

In this regard, we proposed an approach to secure data through a flawless hybrid attribute-based re-encryption scheme.

6.0 Conclusion and Recommendation

6.1 Conclusion

Since data confidentiality is essential in the provision of information security in the cloud, this paper proposes the development of better cryptographic algorithms in the form of a hybrid attribute-based re-encryption scheme to provide for proper encryption methods, because with the growth of computing power application are constantly being developed that can compute the encryption keys being used by the various algorithms currently in use. Also, the research sought to be directed on coming up with systems that will possess self-hardening elements, by use of artificial intelligence, whereby a system will proactively detect and learn a malicious intent to hack it, and on its own, take countermeasures to protect itself by having prior knowledge on various attack profiles and how to

mitigate against such attacks, hence, offering confidentiality to user data. Finally, Cloud Service Providers may consider easy to use steps on how clients can ensure the security of their data in the cloud environment by providing short time training.

References

Coppolino, L., Antonio, S. D., Mazzeo, G., & Romano, L. (2017). Cloud security : Emerging threats and current solutions R. *Computers and Electrical Engineering*, *59*, 126–140. https://doi.org/10.1016/j.compeleceng.2016.03.004

Deng, H., Qin, Z., Wu, Q., Guan, Z., & Zhou, Y. (2020). *Flexible attribute-based proxy re-encryption for efficient data sharing*. *511*, 94–113.

Levitin, G., Xing, L., & Dai, Y. (2018). Co-residence based data vulnerability vs. security in cloud computing system with random server assignment. *European Journal of Operational Research*, *267*(2), 676–686. https://doi.org/10.1016/j.ejor.2017.11.064

P, P. K., P, S. K., & Alphonse, P. J. A. (2018). Journal of Network and Computer Applications Attribute-based encryption in cloud computing : A survey, gap analysis, and future directions. *Journal of Network and Computer Applications*, *108*(June 2017), 37–52. https://doi.org/10.1016/j.jnca.2018.02.009

Rao, B. T. (2016). A Study on Data Storage Security Issues in Cloud Computing. *Procedia - Procedia Computer Science*, *92*, 128–135. https://doi.org/10.1016/j.procs.2016.07.335

Rizvi, S., Ryoo, J., Kissell, J., Aiken, W., & Liu, Y. (2017). A security evaluation framework for cloud security auditing. *The Journal of Supercomputing*. https://doi.org/10.1007/s11227-017-2055-1

Tai, S., & Bermbach, D. (2017). *Cloud Computing*. 1–6. https://doi.org/10.1007/978-1-4614-7163-9

Britt, P. (2019). Cloud Security vs. Security in the Cloud: What's the difference? Similar Sounding Cloud Security has a very different meaning. Accessed on 16.09.2019

Cloud Security Alliance (2016). The Treacherous 12 Cloud Computing Top Threats in 2016. Top Threats Working Group.

Daylami, N. (2016). *Cloud Computing: Demystifying the Elephantine Concept.* California Lutheran University.

Feldman, A. J. (2012). Privacy and integrity in the untrusted cloud. Mountain View, California, USA.Gupta, G., Laxmi, P. R., & Sharma, S. (2014). A Survey on Cloud Security Issues and Techniques. *International Journal on Computational Sciences & Applications (IJCSA). Vol. 4, No, February 2014*

Kepes, B. (2017). Understanding the Cloud Computing Stack: SaaS, PaaS, IaaS. Executive Summary. https://support.rackspace.com/white-paper/understanding-the-cloud-computing-stack-saas-paas-iaas/ [Accessed on 14/07/2019].

National Institute of Standards and Technology (2014). NIST Special Publication 800 – 137: *Information Security Continuous Monitoring for Federal Information Systems and Organizations.* FISSEA 27th Annual Conference

Seema, S. S. & Shaikh, N (2014). *Cloud Computing: Data Separation Issues. International Journal & Magazine of Engineering, Technology, Management, and Research A Monthly Peer Received Open Access International. ISSN No: 2348-4845.*

Samson, T. (2013). 9 Top Threats to Cloud Computing. Infoworld, Feb 25, 2013. http://www.infoworld.com/article/2613560/cloud-security/cloud-security-9-top-threats-to-cloud-somputing-security.html. Accessed on 20.07.2019.