Studying Security of Data in Cloud Computing Through Cryptographic Approach

Hosam F. El-Sofany[1](✉) and Samir A. El-Seoud[2]

[1] Department of Computer Science, Cairo Higher Institute, Cairo, Egypt
hosam_elsofany@hotmail.com
[2] Faculty of Informatics and Computer Science,
British University in Egypt-BUE, Cairo, Egypt
Samir.elseoud@bue.edu.eg

Abstract. Cloud computing is a set of IT services offered to users over the WWW on a rented base. Cloud computing has many advantages such as flexibility, efficiency, scalability, integration, and capital reduction. Moreover, it provides an advanced virtual space for organizations to deploy their applications and to run their systems. For secure communication over Cloud network, data can be protected by the method of encryption. Encryption exchanges that data by any encryption algorithm using the key in twisted form. Only user can access the key used to decrypt the encrypted data. The purpose of encryption is used to preventing leak or secrecy in communications. Encryption algorithms play a huge role in providing data security against bad and malicious attacks. This paper studies the basic concepts and analyzes the essentials of data security issues pertaining to Cloud Computing. Then we elaborate on each issue by discussing its nature. Specifically, we emphasize on issues of protecting data such as: data confidentiality, data integrity, data availability, securing data access, data auditing, enforcing the regulations and compliances regarding to data security and privacy.

Keywords: Cloud computing · Data security · Infrastructure · Cryptography

1 Introduction

From a few years ago, the abstract shapes of cloud are used to denote the internet and cyberspace. Afterwards the cloud has been utilized to represent a more specific area, which is the Cloud Computing. The main idea of cloud computing is to deliver both software and hardware as services. Basically there are three layers of services over the cloud that are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [1]. Individuals and organizations have been considering services over the cloud to cut the costs of expenditure, without any compensation in utilizing recent technologies [2]. A survey conducted by IDC [3] shows the importance of the challenges for those considering cloud computing as an option. It is shown in Fig. 1 that security is the utmost concern. Moving essential data over a network to a third-party resource is not an easy decision to be approved. There should be many guarantees as good performance, availability, and mostly secure transmission

M.E. Auer et al. (eds.), *Interactive Collaborative Learning*, Advances in Intelligent Systems and Computing 545, DOI 10.1007/978-3-319-50340-0_38

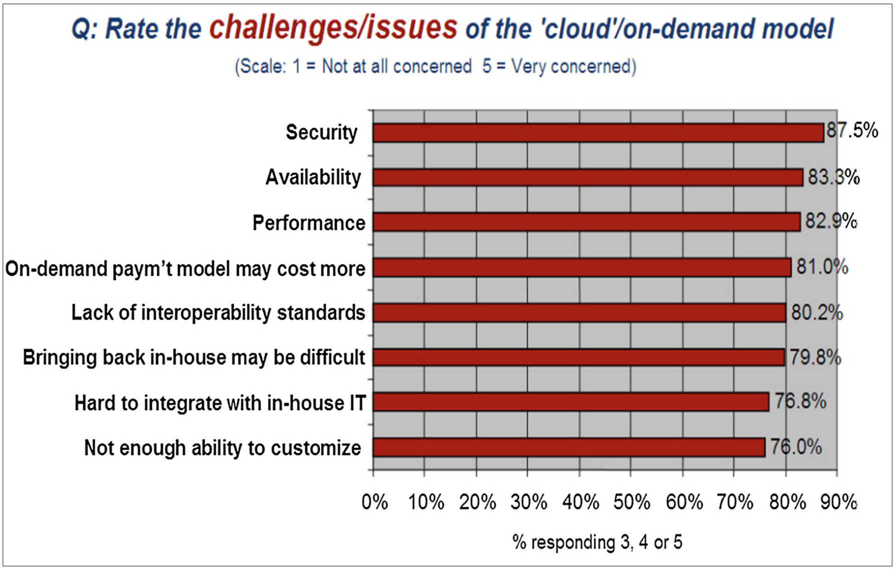

Fig. 1. Challenges in considering cloud computing [3, 4]

and storage. on the other hand, organizations are more reluctant to move important data when the actual infrastructure, precise cost estimation, security level, privacy level, trust, and many other concerns will be unknown.

Cloud Infrastructure as a service provides processing, storage network bandwidth and other fundamental computing resources which allow customers to deploy and run operating systems or applications. For safe communication over public network data can be protected by the method of encryption. Encryption exchanges that data by any encryption algorithm using the key in twisted form. Only user can access the key used to decrypt the encrypted data [5]. The purpose of encryption is used to preventing leak or secrecy in communications [6]. Encryption algorithms play a huge role in providing data security against malicious attacks.

To encrypt the plaintext characters to cipher, their ASCII values are taken and if a character occurs in several places in a plaintext there is a possibility of same cipher text is produced. To overcome the problem, taking ASCII values for the characters to encrypt, preferably different numerals representing the position of ASCII values are taken from magic square and encryption is performed using RSA cryptosystem. This proposed research work provides the security using public key algorithm that ensures the security is improved and also compare these two algorithms and analysis which one is best for encryption in Cloud Computing.

The paper is organized as follows: in Sect. 2, we present an overview about cloud computing and cloud computing security. In Sect. 3, we present a literature review for Cryptography. In Sect. 4 we introduce some previous works for cloud computing security. In Sect. 5, we introduce a Cryptographic approach and analysis results for cloud computing. The paper finally concluded in Sect. 6.

2 Cloud Computing Security

Cloud computing have many advantages in cost reduction, resource sharing, and time saving for new service deployment. While in a cloud computing system, most data and software that users use reside on the Internet, which bring some new challenges for the system, especially security and privacy. Since each application may use resource from multiple servers. The servers are potentially based at multiple locations and the services provided by the cloud may use different infrastructures across organizations. All these characteristics of cloud computing make it complicated to provide security in cloud computing. To ensure adequate security in cloud computing, various security issues, such as authentication, data confidentiality and integrity, and non-repudiation, all need to be taken into account. Currently, Web Services Security (WS-Security) is wildly used in the cloud to provide security for the system. In WS-Security, XML encryption and XML signature are used to provide data confidentiality and integrity. Mutual authentication can be supported by adding X.509 certificate and Kerberos tickets into SOAP message header [8].

As mentioned earlier, there are three types of clouds in general: private cloud, public cloud and hybrid cloud:

1. In a *public cloud*, resources are dynamically provisioned on a fine-grained, self-service basis over the Internet. Services in the cloud are provided by an off-site third-party provider who shares resources and bills on a fine-grained utility computing basis. While in most private clouds, with limited computing resources, it is difficult for a private cloud to provide all services for their users, as some services may more resources than internal cloud can provide. Hybrid cloud is a potential solution for this issue since they can get the computing resources from external cloud computing providers.
2. *Private clouds* have their advantages in corporation governance and offer reliable services, as well as they allow more control than public clouds do. For the security concerns, when a cloud environment is created inside a firewall, it can provide its users with less exposure to Internet security risks. Also in the private cloud, all the services can be accessed through internal connections rather than public Internet connections, which make it easier to use existing security measures and standards. This can make private clouds more appropriate for services with sensitive data that must be protected. While in a hybrid cloud, it includes more than one domain, which will increase the difficulty of security provision, especially key management and mutual authentication.
3. The *hybrid cloud* domains can be heterogeneous networks; hence there may be gaps between these networks and between the different services providers. Even security can be well guaranteed in each of private/public cloud, while in a hybrid cloud with more than one kind of clouds that have different kinds of network conditions and different security policies, how to provide efficient security protection is much more difficult. For example, cross domain authentication can be a problem in a hybrid cloud with different domains. Although some authentication services such as Kerberos can provide multi-domain authentication, but one of the requirements for the multi-domain Kerberos authentication is that the Kerberos server in each

domain needs to share a secret key with servers in other Kerberos domains and every two Kerberos servers need to be registered with each other. The problem here is if there are N Kerberos domains and each of them want to trust each other, then the number of key exchanges is $N(N-1)/2$. For a hybrid cloud with a large number of domains, this will bring a problem for scalability. If different networks in a hybrid cloud using different authentication protocols, this problem can be more complex.

In a cloud, the cloud computing system needs to provide a strong and user-friendly way for users to access all kinds of services in the system. When a user wants to run an application in the cloud, the user is required to provide a digital identity. Normally, this identity is a set of bytes that related to the user. Based on the digital identity, a cloud system can know what right this user has and what the user is allowed to do in the system. Most of cloud platforms include an identity service since identity information is required for most distributed applications. These cloud computing systems will provide a digital identity for every user. For example, user with a Windows Live ID can use cloud computing services provided by Microsoft and user who wants to access cloud computing services from Amazon and Google also needs an Amazon defined identity and Google account. Here, each of these companies is a public cloud.

The problem here is this digital identity can only be used in one private cloud or one public cloud. Users want to access services in the cloud that provided by different clouds will need to have multiple identities, each for one of the cloud. This is obviously not user friendly [7, 8].

3 Cryptography

Cryptography is the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. Cryptography provides us with the tools that underlie most modern security protocols. Nowadays, the most effective method of securing the data is by using cryptographic techniques. Cryptography is the method of storing and transmitting data in form that only those it is intended for can read and process. Basic term s used in cryptography is [17]:

4. The readable data is referred to as PLAINTEXT.
5. The random and unreadable data is referred to as CIPHERTEXT.
6. Process of converting plaintext to cipher text is referred to as ENCRYPTION.
7. Reverse of encryption i.e. Process of converting cipher text to plaintext is known as DECRYPTION.
8. Set of rules dictating how to encrypt and decrypt data are referred to as ALGORITHM (Fig. 2).

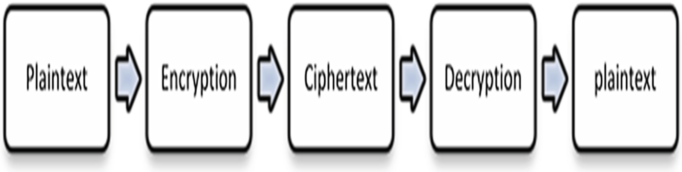

Fig. 2. Cryptography process

- *Cryptosystem*: The hardware or software implementation of cryptography process is termed as cryptosystem. Following services are provided by cryptosystems [17]:
  1. **Confidentiality:** Confidentiality is the need to ensure that information is disclosed only to those who are authorized to view it.
  2. **Integrity:** Integrity is the need to ensure that information has not been changed accidentally or deliberately, and that it is accurate and complete.
  3. **Authentication:** Authentication is the process of confirming correctness of the claimed identity.
  4. **Authorization:** Authorization is the approval, Permission or empowerment for someone to do something.
  5. **Non repudiation:** Non Repudiation is the ability for a system to prove that a specific user and only that user sent a message and it hasn‘t been modified.

- *Public Key Cryptography*: As discussed in RSA data security whitepaper [3] Cryptography uses mathematical algorithms and processes to convert intelligible plaintext into unintelligible cipher text, and vice versa. Applications of cryptography include:
  – Data encryption for confidentiality
  – Digital signatures to provide non-repudiation (accountability) and verify data integrity
  – Certificates for authenticating people, applications and services, and for access control (authorization)

The two main kinds of cryptography are shared secret (*symmetric key encryption*) and *public key* (*Asymmetric key encryption*):

1. *Symmetric Key Encryption:* In symmetric key encryption, encryption key can be calculated from the decryption key and vice versa. With most of the symmetric algorithms, the same key issued for encryption and decryption. The symmetric key is effective only when the key is kept secret by two parties if anyone else discovers the key in any way; it affects both Confidentiality and Authentication. A person with unauthorized symmetric key not only can decrypt messages sent with key but can encrypt new messages and send them on behalf of the legitimate parties using the key.
2. *Asymmetric Key Encryption:* Public key encryption also called as Asymmetric Encryption involves a pair of keys, a public key and a private key, associates with an

entity. Each public key is published, and the corresponding private key is kept secret. Data encrypted with public key can be decrypted only with corresponding private key.

- Public Key Infrastructure (PKI): PKI consists of programs, data formats, procedures, communication protocols, security policies and public key cryptographic mechanisms working in a comprehensive manner to enable a wide range of dispersed people to communicate in a secure and predictable fashion. PKI provides authentication, confidentiality, non-repudiation, and integrity of the message sex changed. PKI is hybrid system of symmetric and asymmetric key algorithms and methods [17].

A public-key infrastructure (PKI) is a framework that provides security services to an organization using public-key cryptography. These services are generally implemented across a networked environment, working conjunction with client-side software, and can be customized by the organization implementing them. An added bonus is that all security services are provided transparently—users do not need to know about public keys, private keys, certificates, or Certification Authorities in order to take advantage of the services provided by a PKI [17].

1. Components of PKI: As shown in [18] there are five components in PKI:
    (a) End Entity: End Entity is a generic term used to denote end-users, devices (e.g., servers, routers), or any other entity that can be identified in the subject field of a public key certificate. End entities typically consume and/or support PKI-related services.
    (b) *Certification Authority (CA*): an entity which issues certificates. One or more in-house servers, or a trusted third party such as VeriSign or GTE, can provide the CA function
    (c) *Registration Authority (RA):* The RA is an optional component that can assume a number of administrative functions from the CA. The RA is often associated with the End Entity registration process, but can assist in a number of other are as well.
    (d) *Repository:* A repository is a generic term used to denote any method for storing certificates and CRLs so that they can be retrieved by End Entities.
    (e) *CRL Issuer:* The CRL Issuer is an optional component that a CA can delegate to publish CRLs.

2. PKI and the Aims of Secure Internet Communication:

The four aims of secure communication on the Internet are as stated earlier: confidentiality, integrity, authentication and non-repudiation. Authentication is the procedure to verify the identity of a user. There are three different factors authentication can be based on. These factors are something the user knows, something the user possesses and something the user is. Something the user knows could be a password that is a shared secret between the user and the verifying party. This is the weakest form of authentication since the password can be stolen through, for example, a dictionary attack or sniffing the network. Something the user possesses could be a physical token like a credit card, a passport or something digital and secret like a private key.

This authentication form is usually combined with something the user knows to form a two-factor authentication. For instance, a credit card and a PIN are something possessed and something known. Something the user is could be something biometric like a fingerprint, DNA or a retinal scan which is unique for the user.

4 Previous Works in Cloud Computing Security

Gartner in [9] recognized seven security risks that are essential to be considered before enterprises make decisions regarding the transformation into a cloud computing model [10]. These problems are as follows:

1. Authorized user access: the potential risk of exposing organizational data over an external processing platform, due to the limited physical, logical and personal controls outside the organizational boundaries.
2. Conformance to regulations: processing data outside the organizational boundaries is still subject to accountability measures, for instance in case of auditing an external third-party space.
3. Storage space: cloud customer has no clue about the exact location of their data that requires service provider commitment to comply with privacy restrictions.
4. Data separation: clouds hold the customers' data over a shared place where data segments are not stored in sequential manner, for that a reliable and well-tested encryption schemes are needed.
5. Recovery: service providers are supposed to make it clear how they will handle disasters and failures.
6. Investigation: breach or intrusion attempts are hard to be tracked and spotted over the cloud due to the dispersion of the data and resources. While in some cases it could be impossible because of the high complexity level.
7. Long-term viability: if a rare case of service provider bankruptcy or acquisition occurs there should be a guarantee of data availability. An organization needs to be sure that it will not lose a huge amount of important data on the long-run.

In [10, 11] the authors examined different security and privacy concerns related to cloud computing. They discussed and outlined the risks, their influences, and the opportunities. Adequate levels of reliability, confidentiality, and sensitive data protection are examples of many security concerns [16]. Clouds as a computing model demonstrate a promising future; at the same time they highly require serious acts to cover their weak points. The weaknesses and problems come from unresolved issues in the existing technologies, which are used to build the cloud. Despite the origins or locations of risks and threats, the cloud security as an issue should be handled in a comprehensive manner [12, 13]. Service providers seek fulfilling security requirements over the clouds, but face different challenges to guarantee high level of security.

The authors in [14] discussed the requirement and challenges, also suggested standardization and management approaches to guide cloud engineers and users. Cloud computing as an approach introduces new risks, influences others, and magnifies some. These risks and their effect on security risks and vulnerabilities were explained in [11]. Standardizing the cloud services security is an important issue that emerged due to the increased demand and importance of clouds [15]. For instance, standardized Security Level Agreement (SLA) guarantees transparent assurance and increases the trust among cloud adopters. These standardized guarantees assist in having mutual trust, reduced risks, and better dissemination of cloud service among organizations as customers, service providers and investors.

5 Cryptographic Approach and Analysis Results for Cloud Computing

As shown in the previous sections, "Cryptography" is the study of mathematical techniques related to characteristic of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. The two famous encryption techniques that can be implemented in the cloud computing security systems are RSA and Magic Square approaches:

- RSA: this algorithm is developed by Rivest, Shamir and Adleman. In the asymmetric method, *private key* and *public key* are used. Both the *sende*r and *receiver* know the public key and it is used to encrypt the data. The owner of the private key can only decrypt the message. The public and private keys are always in pairs, it is difficult to derive at the private key from the public key which is shared. That is why this method is considered to be more secure than the symmetric method. RSA adopted public key cryptography algorithm. The encryption process can be done either at the customer's end or at the service provider's or at the vendor's end. The customer can do the encryption process and it will increase the computation time and usage which in turn increases the cost. Furthermore there may be no guarantee for the proper implementation of the encryption process. Because of these, it is appropriate to do the encryption process at the Client's end [18].
- Magic Square: magic square is a square array of numbers consisting of the distinct positive integers 1, 2, …, n^2 arranged such that the sum of the n numbers in any horizontal, vertical, or main diagonal line is always the same number. Given an $n \times n$ matrix of the integer 1 to n^2 such that the sum of every row, column and diagonal is the same. Then n rows the sum of all the numbers in the magic square must be (n*M). But the numbers being added are 1, 2, …, n^2, and so $1 + 2 + 3 + \ldots + n^2 = n*M$, i.e., $\sum_{i=1}^{n^2} \mathrm{i} = n * M = \frac{n^2(n^2+1)}{2}$. Then solving for M gives $M = \frac{\mathrm{n}(n^2+1)}{2}$. Thus, $a_{3\times 3}$ normal magic square must have its rows, columns and diagonals adding to $M = \frac{3(3^2+1)}{2} = 15$, and in case of $a_{4\times 4}$ we get $M = \frac{4(4^2+1)}{2} = 34$. The magic sum for an $n \times n$ normal magic square can be found by filling the $n \times n$ square with the numbers 1, 2, …, n^2 first going across the top row, then the second row, and so on and then adding the numbers along either of the diagonals. There are three types of magic squares:
  - The odd order magic square is referred to M is an odd number $M = 2n + 1$ where $n = 0, 1, 2, 3\ldots$
  - The even order magic square is referred to M is an even number divisible by both 2 and 4 $M = 4(n + 1)$ where $n = 0, 1, 2, 3\ldots$

– The singly even order magic square are referred to M is an even number divisible by 2 but not by 4 $M = 2(n + 3)$ where $n = 0, 1, 2, 3\ldots$

So we can construct a set of different Magic Square of order $n \times n$; $16 \geq n \geq 8$ and n is even as far as possible and each magic square corresponds to one ASCII set. To encrypt the character, use the ASCII value of the character to determine the numeral in the magic square by considering the position in it. Let NP and NC denote the numeral of the plaintext and cipher text respectively. Based on NP and NC values, all plaintext and cipher text characters are encrypted and decrypted respectively.

5.1 Security Analysis Results

1. The security analysis consists of analyzing various security properties such as *Data Confidentiality*, *Authentication* and *Integrity of the data*.

   - *Data Confidentiality* is analyzed by comparing it with various data Encryption algorithms such Advanced Encryption Standard (AES) or Data Encryption Standard (DES) which uses the symmetric key for encrypting the data.
   - *Authentication*: A new user is added and tries to access the data over a cloud. Authentication is performed with the help of the password set by the user during registration.
   - *Integrity*: Ensures that the data integrity is maintained and the data over the cloud is secured.

2. RSA algorithm analysis

The RSA algorithm is implemented with different input files of different sizes, through local and cloud environments, as shown in Table 1.

Table 1. Comparison between processing times in local and Cloud sites

Size of input file	RSA in local site	RSA in cloud
3 MB	690.3	390.2
6 MB	789.2	402.6
12 MB	801.4	420.9
24 MB	923.4	480.2

3. RSA with Magic Square approach

The time taken for encryption and encryption of different files with different sizes in simulated and parallel environment using RSA public key crypto system with magic square is shown in Table 2.

Table 2. RSA Encryption and Decryption time

Size of input file	Encryption time (ms)	Encryption time (ms)	Total Time
1 MB	225	250	475
2 MB	435	465	900
4 MB	896	910	1806
8 MB	1760	1845	3605

6 Conclusions

In this paper, we will present the basic concepts behind the cloud and introduce the security issues underlying cloud computing. The study focuses on a set of crucial security issues pertaining to storing and accessing data in cloud computing such as: data confidentiality, data integrity, data availability, securing data access, data auditing, enforcing the regulations and compliances regarding to data security and privacy. We will analyze each issue and discuss the possible solutions based on existing techniques.

References

1. Armbrust, M., Fox, A., Grith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: A view of cloud computing. Commun. ACM **53**(4), 50–58 (2010). http://dx.doi.org/10.1145/1721654.1721672
2. Jensen, M., Schwenk, J., Gruschka, N., Iacono, L.: On technical security issues in cloud computing. In: IEEE International Conference on Cloud Computing, Bangalore, pp. 109–116, 21–25 September 2009
3. Gens, F.: New IDC It Cloud Services Survey: Top Benefits and Challenges (2009)
4. http://blogs.idc.com/ie/?p=730
5. Cloud Security Alliance (2010). Top threats to cloudcomputing, version1.0. http://www.cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf
6. AnoopMS, Publickey Cryptography (Applications Algorithm and Mathematical Explanations)
7. Yan, L., Rong, C., Zhao, G.: Strengthen cloud computing security with federal identity management using hierarchical identity-based cryptography. In: Jaatun, M.G., Zhao, G., Rong, C. (eds.) CloudCom 2009. LNCS, vol. 5931, pp. 167–177. Springer, Heidelberg (2009). doi:10.1007/978-3-642-10665-1_15
8. Chappell, D.: A Short Introduction to Cloud Platforms. http://www.davidchappell.com/CloudPlatforms–Chappell.pdf
9. Brodkin, J.: Gartner: Seven Cloud-Computing Security Risks, InfoWorld (2008). http://www.infoworld.com/d/security-central/gartner-seven-cloud-computing-security-risks-853
10. Chen, D., Zhao, H.: Data security and privacy protection issues in cloud computing. In: International Conference on Computer Science and Electronics Engineering, vol. 1, Hangzhou, pp. 647–651, 23–25 March 2012
11. Grobauer, B., Walloschek, T., Stocker, E.: Understanding cloud computing vulnerabilities. IEEE Secur. Priv. **9**(2), 50–57 (2011). http://dx.doi.org/10.1109/MSP.2010.115
12. Almorsy, M., Grundy, J., Müller, I.: An analysis of the cloud computing security problem. In: Proceedings of the 2010 Asia Pacific Cloud Workshop, Australia, 30 November 2010

13. Leavitt, N.: Is cloud computing really ready for prime time? Computer **42**(1), 15–20 (2009). http://dx.doi.org/10.1109/MC.2009.20
14. Popović, K., Hocenski, Z.: Cloud computing security issues and challenges. In: Proceedings of the 33rd International Convention in MIPRO, pp. 344–349 (2010)
15. Ramgovind, S., Elo, M., Smith, E.: The management of security in cloud computing. In: Information Security for South Africa, Sandton, pp. 1–7, 2–4 August 2010
16. Kandukuri, B.R., Paturi, V.R., Rakshit, A.: Cloud security issues. In: Proceedings of the 2009 IEEE International Conference on Services Computing, Washington DC, pp. 517–520. 21–25 September 2009. http://dx.doi.org/10.1109/SCC.2009.84
17. Heena, K., Chouhan, D.S.: Building trust in cloud using public key infrastructure- a step towards cloud trust. Int. J. Adv. Comput. Sci. Appl. (IJACSA), **3**(3) (2012)
18. Rivest, R., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**(2), 120–126 (1978)