Asian Journal of Information Technology 13 (11): 707-714, 2014
ISSN: 1682-3915

Proposal for a New Authentication and Encryption Algorithm for Data at Rest on Cloud Environments

Sreevidya Subramanian and Ananthi Seshasaayee
Department of Computer Applications, Vels University, Chennai, India

Abstract: Cloud computing is technology in this modern era for handling enterprise IT in a more effective manner. It is providing a means for business to meet their IT demands seamlessly simply by paying for the usage as needed and on demand service delivery opportunities. The term cloud computing is continuously evolving with each industrialist giving their own explanation. It is true that no new technology comes without hindrances but with the due efforts from technology experts and SMEs all of these pain points will wash away. Despite of all the positive features it provides, not all industries are effectively taking the decision to move towards cloud environment primarily due to its painful security issues. The current problem now with IT organizations is that without assessing their current risk levels, they migrate on to public cloud infrastructure without thinking through the security of confidential data that is going to be put through in to the cloud. This leads in to malicious insider attacks, hacking confidential data either at transit or at rest. In this study, we will look in to the security issues in the field of cloud computing and address this by providing a stable/fool proof approach and model for session oriented encryption based on unicode storage concept for data at rest. The proposed model will also provide a framework for secure logging and encryption for data at rest.

Key words: IaaS, virtualization, encryption, software as a service, data centers, utility computing, UTF-8, SaaS, EaaS

INTRODUCTION

Cloud computing may look like a new technology but this has been evolving from many distributed technologies. It is just a new name for an old fabric. It originated with the concept of virtualization which led to reduction of total cost of ownership for IT Industry by taking the virtualization to the next level. Cloud technology now helps businesses to reduce their global footprint and moving towards Greener ecosystem. Cloud is an end result of convergence of two technologies (Huang *et al*., 2012) which is clearly explained with Fig. 1. Cloud computing has the following characteristics (Reese, 2009; Buyya *et al*., 2009):

Virtual estate: Sharing of resources and optimal scalability by virtualized platforms.

Dynamic provisioning: On demand provisioning of Servers, resources with a reliable yet faster turn-around time.

Speed and flexibility: Many organizations today can agree to the fact that ICT cannot keep up with the pace and agility demand by business, instead of providing the required support to business, ICT has rather become a bottleneck. Most companies are ready for the challenging changes of future by moving on to cloud based environment, the most promising feature/benefit of cloud is rather the speed and flexibility it provides.

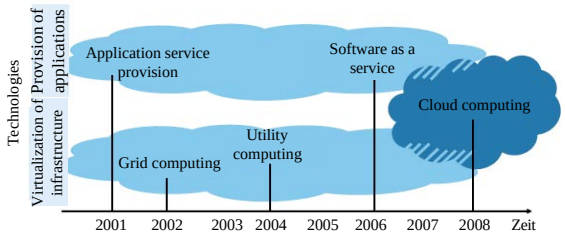

Fig. 1: Two trends converging resulting in to cloud computing

Reliability: This is improved through the use of multiple redundant sites, making cloud computing infrastructure suitable for DR and HA purposes.

Sustainability: This comes about through improved resource utilization, more efficient systems and carbon neutrality.

Corresponding Author: Sreevidya Subramanian, Department of Computer Applications, Vels University, Chennai, India

Cost effectiveness: Studies show that large companies can also make significant savings through cloud computing. The exact scale of these savings depends on the concrete customer situation. A particularly deciding influence on the cost reductions is the question of how strongly virtualized the resources were before the use of cloud technologies.

CLOUD SERVICE MODELS

Three core options comprise the service model within the cloud computing environment. However, in this era of new age methodologies and process standards, many more of such service models are evolving namely DaaS (Data as a Service), BaaS (Backup as a Service), RaaS (Risk Assessment as a Service), AaaS (Audit as a Service), SecaaS (Security as a Service), EaaS (Encryption as a Service), etc.

Software as a Service (SaaS) comprises end user applications delivered as a service, rather than as traditional, on-premises software. The most commonly referenced example of SaaS is Salesforce.com which provides a Customer Relationship Management (CRM) System accessible via the Internet.

Platform as a Service (PaaS) provides an application platform or middleware as a service on which developers can build and deploy custom applications. Common solutions provided in this tier range from APIs and tools to database and business process management systems, to security integration, allowing developers to build applications and run them on the infrastructure that the cloud vendor owns and maintains. Microsoft's Windows Azure platform services are often referenced as PaaS solutions at this middleware tier.

Infrastructure as a Service (IaaS) primarily encompasses the hardware and technology for computing power, storage, operating systems or other infrastructure, delivered as off-premises, on-demand services rather than as dedicated, on-site resources such as the Amazon Elastic Compute Cloud (Amazon EC2) or Amazon Simple Storage Service (Amazon S3).

CONTROL SEGREGATION IN CLOUD

Extending an organization's access management services into the cloud is ad hoc, dependent on your cloud service provider and rudimentary at best. However, with that said, there are some steps you can take to utilize cloud services as they become more mainstream:

- Ask your cloud service provider to support open standards for access management
- Standardize and automate your user provisioning as much as possible
- Create a centralized entitlement management mechanism within your organization
- Extend policies and procedures regarding access management to include cloud services

At the object level, the common control segregation is explained with the help (Fig. 2).

CLOUD DEPLOYMENT MODELS

Four deployment models have been identified for cloud architecture solutions (Mell and Grance, 2011) described below:

Private cloud: The cloud infrastructure is operated for a private organization. It may be managed by the organization or a third party and may exist on premise or off premise.

Community cloud: The cloud infrastructure is shared by several organizations and supports a specific community that has communal concerns (e.g., mission, security requirements, policy and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

Public cloud: The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

Hybrid cloud: The cloud infrastructure is a composition of two ormore clouds (private, community or public) that remain uniqueentities but are bound together by standardized or proprietary technology, that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

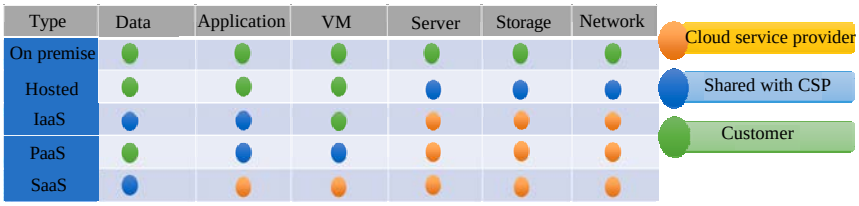

Fig. 2: Segregation of control between customer, CSP and shared model

CURRENT SECURITY ISSUES

Cloud computing due to its features and design characteristics may provide a host of benefits which may include centralization of security, redundancy and HADR capabilities. While many traditional risks are countered effectively, the most of the incidents registered on cloud environment are a whopping number amounting to 172.

Cloud computing has unique attributes that require risk assessment in areas such as availability and reliability issues, data integrity, recovery and privacy and auditing. Based on the above statistics of recorded incidents, we also have the top threat identified for cloud in 2011 (Fig. 3).

Figure 4 shows the frequency of occurrence of the existing seven CSA threats and five new threats proposed by CSA Expert group. The three most frequent incidents are:

- T2: "Insecure interfaces and APIs" with 51 incidents accounting for 29% of all threats
- T5: "Data loss and leakage" with 43 incidents accounting for 25% of all threats reported
- T8: "Hardware failure" with 18 incidents accounts for 10% of all threats reported

All other threats have 15 or fewer cloud vulnerability incidents each, accounting for 8.5% or less. At the outset,

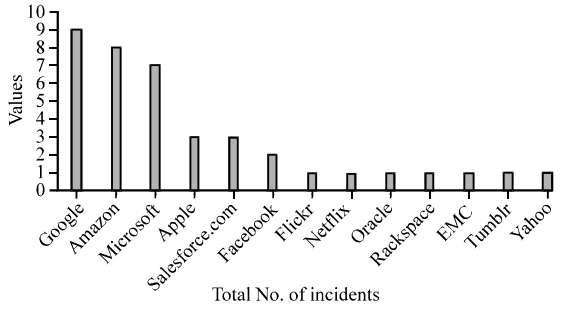

Fig. 3: Total No. of incidents reported by CSPs

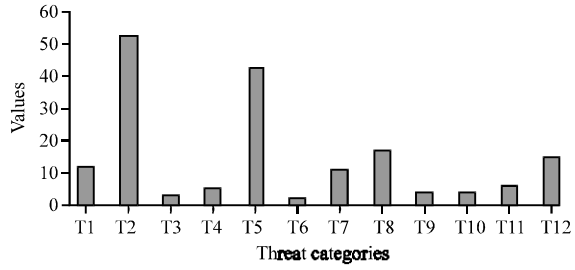

Fig. 4: Total No. of incidents reported per threat category

one must identify what needs to be protected against security breaches. Is it only data and software that needs protection? Well, to a majority extent this could be true as these are the key vulnerabilities today. Cloud means sharing of resources as well which implicit that the activity performed by a user might also be visible to the other leading to side channel or man in the middle attacks.

When using shared resource for business critical activities or transactions, it will be very difficult to track malicious and unethical activities. There are numerous tools to identify the risk in a proactive way, we are still not able to eradicate the threats that are increasing day by day.

Based on the above statistics, it is very clear that the two most demanding threats for us would be to have a streamlined use of secured APIs and to minimize data loss/leakages. Common issues around data loss/leakages are around encryption of critical data when transmitted or stored in cloud especially when key management and key generation lies within the purview of the cloud service providers and its corresponding authentication framework used.

Delegating data control to the cloud directly leads to an increase in the risk of data compromise as the data becomes accessible to unwanted parties and thus posing a huge risk to confidentiality. A number of concerns emerge regarding the issues of multitenancy, data remanence, application security and privacy (Blum *et al*., 1990).

A new technique is used today where by the attackers will send a lot of fake cipher text messages to the target server and they retrieve error messages sent out by that server to interpret the actual text that is communicated. This is now being quite popular which poses a huge threat to any encryption performed through the internet. To minimize the unknown threats that are surrounded by cloud and to start looking at base-lining a strong encryption and authentication framework, unicode mapping and color code encryption technology is being proposed along with a Session based authentication framework.

ENCRYPTION AND ITS VULNERABILITIES

Encryption is a process by which the data that needs to be shared with one or more parties is converted in to a form that cannot be easily understood or interpreted by unauthorized people. The converted text is called ciphertext.

This technology has its roots from ancient techniques of communication. The cipher generation can

be very simple or complex depending on the type of data transferred and the type of vulnerability expected during transmission.

Both encryption and decryption process is dependent on a common factor called a key which determines the uniqueness of the text being encrypted. Encryption has gained a lot of importance in wireless communications. This is because of the foreseeable vulnerability that wireless circuits are easier to tap than their hard-wired counterparts (Chow *et al*., 2009).

Based on the strength of encryption, i.e., the harder it is for the ciphertext to be broken by unauthorized attackers, the better it is. But the cost increases proportionately to its strength. Attributes of framing a secure and strong encryption scheme are described as follows:

- Strong Authentication Model
- Level of privacy and confidentiality
- Integrity of encryption schemes
- Non-repudiation
- Reliability and availability

SECURITY REQUIREMENTS FOR DATA STORAGE ON CLOUD

According to the literature (ENISA, 2009; Francis *et al*., 2001; Furht and Escalante, 2010; Gellman, 2009; Kamara and Lauter, 2010; Khorshed *et al*., 2012; Krutz and Vines, 2010; Ho, 2009), security is the major concern for the individuals and organizations when moving into the cloud. One of the major challenges that face cloud computing is how tosecure the data when stored in the cloud. Each time an individual, an organizationor a government adopts the cloud to store data, privacy or confidentiality questions may arise. The main idea is whether to trust the cloud provider or not. As we will see in this thesis, many of today's cloud providers claim that they provide a high level of security and protection to their clients data. However, all these are just claims written in the Service Level Agreement (SLA) but with no enforcement.

In other words, the users need to trust the cloud provider in order to get the benefit of utilizing cloud technologies.

PROPOSED SBA ENCRYPTION FRAMEWORK

This research work identifies the potential pitfalls of data security and tries to propose a framework for enabling secured data storage through session based authentication model coupled with combined unicode and color code mapping technique.

Cryptographic algorithms are variety in number for the framework above, we are using one of the symmetric cipher. Symmetric cipher is divided in to two broad categories namely block cipher and stream cipher. The design will be based on stream ciphers. Stream ciphers encrypt bits individually. This is achieved by adding a bit froma key stream to a plaintext bit. There are synchronous stream ciphers where the key stream depends only on the key and asynchronous ones where the key stream also depends on the cipher text.

CURRENT MODEL IN USE WITH CLOUD

Data owners will out source encrypted data on to the cloud. For privacy issues encrypting the data seems to be abetter choice. However, there exist a lot of confusions and issues over key management when it comes to cloud. Who will own the key management? Will it be with the cloud provider or with the customer? What will be the win-win situation for both parties without the risk of losing data? Current setup of cloud environment is explained with Fig. 5.

Problem statement: Generically, data on cloud is classified into two types; data in process and data in storage. There is a need to devise a proper encryption strategy along with the right access control mechanism for encrypting data on to the cloud. This depends on a variety of factors:

- Depends heavily on encrypted data in the cloud environment
- Protecting the data from the cloud itself especially for data at rest rather than data at transit is of prime importance to business. This is due to the fact that customers are unable to keep track of what is happening to their data

However, encrypted data on the cloud places limitationsupon data searches and queries.

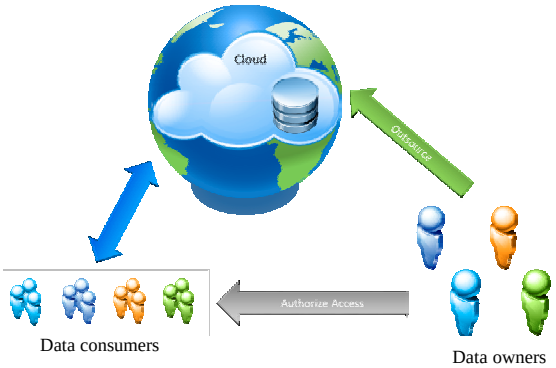

Fig. 5: Current model in use with the cloud

CLOUD FRAMEWORK

The design base for the crypto system is perfectly secure under the following conditions:

- The key k used is unique for each session and no two subsequent session will have the same key for authentication and encryption
- The size of the key should be almost equivalent to that of the size of plain text
- The resultant cipher text will be mapped to respective unicode for storage on to the cloud. This enhances dual security of the data due to the fact that the resultant data cannot be displayed when hacked
- The design is dedicated for encryption of data at rest and not for data at transit. When the same set of cipher text/encrypted data is selected for transit then another conversion is required. This builds additional security

Key industries like banks/financial organizations and those who have high risk confidential data never choose cloud as a viable option primarily because they are not sure about who controls the data in the cloud is it the user or the cloud provider? It is true that the cloud customers have theoption to trade the privacy of their data for the convenience of software services (e.g., web based email and calendars). However, this is generally not applicable inthe case of government organizations and commercial enterprises (Marks and Lozano, 2009; Rittinghouse and Ransome, 2010; Roiter, 2009; Bellare *et al*., 1998). Such organizations will not consider cloud computing as a viable solution for their ICT needs, unless they can be assured that their data will be protected at least to the same degree that in-house computing offers currently.

MEASUREMENT CRITERIA OF THE NEW MODEL

The proposed SBA encryption framework designed for use with cloud for data at rest will be measured against the following key parameters:

Data confidentiality: This refers to the prevention of intentional or unintentional unauthorized disclosure of information. This means that we need to ensure only authorized clients with appropriate access control can operate on to the cloud application.

Data integrity: This ensures that the stored data has not been inappropriately modified (whether accidentally or deliberately). Data integrity becomes more challenging when adopting cloud computing where cloud customers outsource their dataand have no (or very limited) control over their stored data from being modified by the storage service provider. Thus, cloud customers are aiming to detect any unauthorized modification of their data by the cloud storage provider.

Data availability: Ensures that users are able to obtain their data from the cloud provider when they need it. Cloud customers want to be sure that their data is always available at the cloud storage.

Design principle: The design will be based on the following principles:

- The key stream s_0, s_1, s_2, ... is generated by a color palette which will form the key for that particular session
- The key stream is only known to the legitimate communicating parties and for encryption and decryption of files in that particular session
- Every key stream bit s_i is only used once is somewhat nearly same as that of one time pad

PROBLEM DEFINITION

The purpose of an Encryption algorithm is to protect the secrecy of messages which are sent over an insecure channel. Let E be an encryption function which converts a given plain text in to a cipher text using a code factor as a key phrase. The encrypted cipher text is strong and efficient for use with data at rest.

METHODOLOGY

The proposed algorithm consists of the following mathematical transformation functions:

- User registers using an application (frontend) and he chooses a password for his login
- This must get stored in a user table for which a corresponding session table is created. The primary key of user table will then be the foreign key in session table
- Whenever he/she wishes to use that application, he/she must login and establish a session. During such logon procedure, the user is asked to select a color code from the list of color palette displayed for him/her
- The corresponding session table has to be updated with the session id, user details and the selected color code of that session
- An encryption function E and a decryption function D = E-1

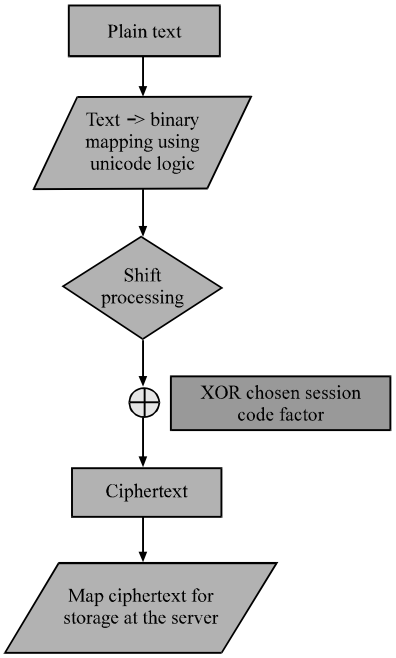

Fig. 6: Encryption process flowchart

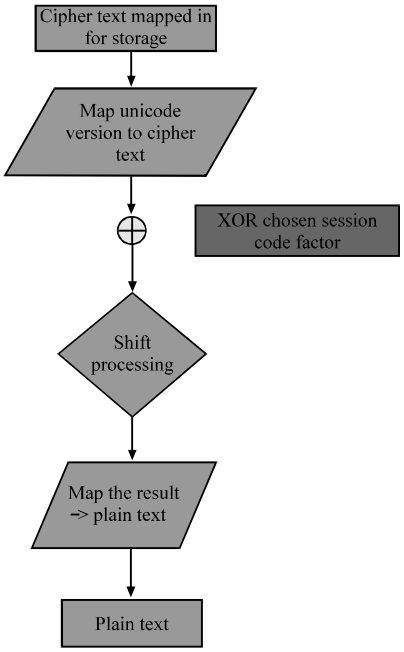

Fig. 7: Decryption process flowchart

We can generate 16.7 million different color combinations here as unique key. This ensures that we randomly generate color codes and that will serve as a viable key for us. The complete process flow is depicted in the below representation.

Encryption process:

- Convert plain text to unicode equivalent
- Map unicode in to binary (8 bits)
- User chooses a code factor when he/she first registers in to the application, this will become their inherent key for encryption
- Performing encryption of data at rest specific to their session
- When the encryption is to be done, the mapped binary text will go through shifting process
- The resultant cipher will then be XOR'd with code factor to produce cipher text (Fig. 6)

Decryption process:

- The cipher text will first be XOR'd with code factor to produce an intermediate text
- The resultant intermediate text will go through the shifting process
- This will result in the binary text (decrypted form)
- The resultant binary text will be mapped to its corresponding unicode
- The unicode is then mapped to ASCII characters (Fig. 7)

IMPLEMENTATION

The methodology of implementing the proposed prototype is divided in to the following phases:

Planning: During this step the requirements design and specification is prepared. The functionalities that are already implemented for various encryption schemes will be analyzed and evaluated against cloud environment suitability. General algorithmic design will be planned on paper.

Exploring: During this phase, the first step of developing this prototype namely encryption process inputs and outputs are generated. The basic encryption process will be tested. Similarly a decryption prototypic on its simplistic form will be generated and tested.

Coding: Based on the above prototype, the design will further be elevated and refined to include the encryption standard of the proposed concept. Coding will continue to incorporate the features specified in the objectives.

Testing: In this phase, the software will be tested with the sample test case scenarios. All the features related to this model will be tested in this phase. Further to this the performance efficiency will also be measured.

Documentation: During this phase, the actual prototype development specification will be document along with the testing scenarios.

Table 1: Selective measurement of parameters

Key parameters	Observations
Data confidentiality	The proposed model will ensure that only authorized users will be able to access the cloud application. In addition to this, the user cannot select two unique session keys consecutively, there will be validation done in order to ensure that no two sessions can have the same unique code factor
Data integrity	The prescribed encryption standard imposes a unique positive aspect of data integrity, i.e., key management, the key management here is done by the application itself and it cannot be controlled by the cloud service provider. With the advent of Bring your own encryption standard in the market, the customers have the provision now to get their encryption standard enabled in the cloud application as the customer will own the application and its data. This encryption concentrates primarily for data at rest but when it is extended to data at transit, it can also serve as a dual encryption saving the integrity of data
Data availability	This framework ensures that both users and customers have a hold on their cloud storage. With the ability of data replication techniques and DR strategy efficiencies available in the market, additional backup of sensitive data on to customer provided storage is also a possibility

The proposed model concentrates on encryption of data at rest so that any hacking will not be able to interpret the data easily and it is more secure.

In order to transfer the encrypted data over emails or to display, it's mandatory to store the resultant cipher text as byte arrays only. This ensures secure transfer and stable cipher text.

MEASUREMENT OF PARAMETERS

Based on the test output it is evident that the framework not only safeguards data but also proves that it is effective for the below selective measurement parameters (Table 1).

CONCLUSION

The conclusions that can be drawn is that the security proposal when implemented as part of future scope of research will function very well and provide good security together with an ease of use for clients that do not have so much technical knowledge.

In the field of security data hiding is the most important task. Cost of the security and efficiency will depend on the confidentiality and sensitivity of the data. So, this type of data hiding, proposed model will be more secure. Some data strings, e.g., password sending, a small information and costly data requires very much security and will be able to leverage proposed model. When security, efficiency and cost is prime concerned rather others parameters then proposed model will be best suitable for use. This is ideally not the public-private key based encryption where in the public key is supposed to be distributed to all users by the cloud service providers. However, this model will be proving useful when combined with multi factor authentication where user's credentials are well validated before the user is even allowed to access the data and manipulate them. Further researches digging deep in to this model can be used to increase the capability of data hiding and correspondingly a level of security can be increase.

RECOMMENDATIONS

This model can be implemented in the cloud based applications where data is stored on the cloud. The main criteria and parameters considered in this model can be increased to develop an extended model. Further research work on this model is planned and will continue to be improvised.

REFERENCES

Bellare, M., A. Desai, D. Pointcheval and P. Rogaway, 1998. Relations among notions of security for public-key encryption schemes. Proceedings of the 18th Annual International Cryptology Conference, August 23-27, 1998, California, USA., pp: 26-45.

Blum, M., P. Feldman and S. Micali, 1990. Proving security against chosen ciphertext attacks. Proceedings of thel 8th Annual International Cryptology Conference, August 23-27, 1998, California, USA., pp: 256-268.

Buyya, R., C.S. Yeo, S. Venugopal, J. Broberg and I. Brandic, 2009. Cloud computing and emerging IT platforms: Vision, hype and reality for delivering computing as the 5th utility. Future Gener. Comput. Syst., 25: 599-616.

Chow, R., P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka and J. Molina, 2009. Controlling data in the cloud: Outsourcing computation without outsourcing control. Proceedings of the ACM Workshop on Cloud Computing Security, November 13, 2009, Chicago, Illinois, USA., pp: 85-90.

ENISA., 2009. Cloud computing information assurance framework. European Network and information Security Agency. https://www.enisa.europa.eu/activities/risk-management/files/deliverables/cloud-computing-information-assurance-framework.

Francis, P., S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt and L. Zhang, 2001. IDMaps: A global internet host distance estimation service. IEEE/ACM Trans. Networking, 9: 525-540.

Furht, B. and A. Escalante, 2010. Handbook of Cloud Computing. Springer, USA., ISBN: 9781441965240, Pages: 653.

Gellman, R., 2009. Privacy in the clouds: Risks to privacy and confidentiality from cloud computing. The World Privacy Forum, pp: 1-26. http://www.worldprivacyforum.org/www/wprivacyforum/pdf/WPF_Cloud_Privacy_Report.pdf.

Ho, A.D., 2009. Cloud strikes all the right chords but security concerns keep it from hitting the perfect pitch. November 2009. http://www.ap.idc.asia/.

Huang, I., R. Guo, H. Xie and Z. Wu, 2012. The Convergence of Information and Communication Technologies Gains Momentum. In: The Global Information Technology Report 2012: Living in a Hyperconnected World, Dutta, S. and B. Bilbao-Osorio (Eds.). Chapter 1.2, The World Economic Forum and Insead, Geneva, pp: 35-46.

Kamara, S. and K. Lauter, 2010. Cryptographic Cloud Storage. In: Financial Cryptography and Data Security, Sion, R., R. Curtmola, S. Dietrich, A. Kiayias, J.M. Miret, K. Sako and F. Sebe (Eds.). Springer, New York, pp: 136-149.

Khorshed, M.T., A.B.M. Ali and S.A. Wasimi, 2012. A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. Future Generation Comput. Syst., 28: 833-851.

Krutz, R.L. and R.D. Vines, 2010. Cloud Security: A Comprehensive Guide to Secure Cloud Computing. Wiley, New York, ISBN: 9780470921449, Pages: 504.

Marks, E.A. and B. Lozano, 2009. Executive's Guide to Cloud Computing. Wiley, New York.

Mell, P. and T. Grance, 2011. NIST definition of cloud computing. Special Publication 800-145, Technical Report, National Institute of Standards and Technology, Gaithersburg, MD., USA.

Reese, G., 2009. Cloud Application Architectures: Building Applications and Infrastructure in the Cloud (Theory in Practice). O'Reilly Media, Canada.

Rittinghouse, J. and J. Ransome, 2010. Cloud Computing: Implementation, Management and Security. CRC Press, Florida, USA.

Roiter, N., 2009. How to secure cloud computing. Information Security Magazine, June 21, 2009. http://searchsecurity.techtarget.com/magazineContent/How-to-Secure-Cloud-Computing.