A Study on the Risks and Mitigation Strategies Related to Data Security in Cloud Computing

S R Phanindra Kumar Narayanasetty[1] and Srikar Madhira[2]

1. M.Sc (S.E), BTH, Sweden

2. M.Sc (S.E), BTH, Sweden

Abstract: The main goal of this paper is to identify the risksand mitigation strategies that are related to data security in cloud computing. This paper tries to provide two checklists: one for the risks and the other for the mitigation strategies. The presented checklists are validated with the help of a survey and then, the top five or the most prevalent risks, which affect the data security hugely, are identified with the help of the survey. The top five or the mostly used mitigation strategies to counter the risks are also identified with the help of the survey. These identified top five risks and mitigations strategies are presented in a separate table. The threats that might be encountered with the research and conclusions with future work are also discussed.

Keywords*: Cloud Computing, Challenges, Risks, Mitigation Strategies, Checklist.*

I. INTRODUCTION

Cloud computing is one of the most rapidly and promisingly growing parts in software and IT industry. Cloud computing is one of the effective and efficient technologies as it can provide the benefits like reduction in costs, increased flexibility and its ability to adapt to the demands [23]. This technology has advanced on its previous versions like grid and utility computing and is expanding rapidly and multiplying itself in industries [22]. The storage in a cloud enables the users to store their data in the cloud and have an access to the applications and all without the burden of hardware and software management [20]. Basically, cloud is an internet based technology. The word cloud computing refers to the applications which are present in the cloud as a service on the internet as well as the hardware or software which is used in order to provide the applications [23]. The benefits provided by cloud also include more scalability, services which are better fault tolerant and increased performance, higher computing resources, elimination of the need to plan on hardware provisioning and many others [1]. The benefits promised by this technology have attracted many industries which have Shifted to this type of computing. Some of the prominent industries and web services which practice cloud computing are amazon, salesforce.com, Google, IBM, Microsoft, Sun, HP and Oracle [21].

But in spite of the benefits promised, the cloud computing technology faces many challenges and problems. The main problem is related to the data security. Data security refers to the security or safety of the data which is stored inside the cloud and on remote machines that are not personal [22]. Many users of cloud computing state that data, like their business information and critical IT resources which are present in the cloud are likely to be attacked and their security is the primary concern [1]. Also, security is a key concern, as the aspects like integrity, authenticity, confidentiality,

auditability of data, tools, transactions and so on are the key features for any business to remain competitive and functional [2]. Therefore, there is a need to ensure the efficient security of the data that is present in the cloud. There are various mitigation strategies in order to ensure security for the data inside the cloud. But we could see that, there was not much effort done in order to present the risks in data security, the mitigation strategies in an effective manner.

The main goal of this research is to identify the risks and mitigation strategies related to data security in cloud computing. The risks and mitigation strategies are tabulated as two separate checklists. These checklists are validated with those in the industry with the help of a survey. After this, the list of top five or the most prevalent risks and the most efficient mitigation strategies from the checklist are identified and presented in a separate table.

For this research, we have chosen to formulate two checklists. One of the checklists consists of the risks that are related to data security in cloud computing. The second checklist consists of the mitigation strategies that are used to counter the risks. The identification of risks and mitigation strategies are done with the help of a systematic literature review. After the identification, they are formulated as two separate checklists. Then, a survey is conducted in order to validate the checklists and to collect the top five risks and mitigation strategies.

The following sections contain the research definition and plan, research methodology, research operation, analysis, discussion and conclusions.

II. MOTIVATION

As mentioned in the previous section, security is an important aspect in the cloud computing area. If the problems or the risks are not identified properly, this might lead to critical losses to the industry or the loss of valuable data. If the mitigation strategies for the risks are also not identified properly, there might be a danger that the industries may be unable to prevent the risks that may occur. Therefore, the necessity to identify the risks that are prevalent in data security motivated us in choosing the topic. We could see that there was not much effort done in order to present the risks and mitigation strategies in an effective manner. The authors [1][2][3][4][5][13][14][15][16][17][18] have spoken about some of the risks or mitigation strategies, but haven't put in much effort in presenting them in a user friendly way.

Therefore, we have chosen to identify the risks and mitigation strategies and present them in the form of a checklist. We also tried to identify the most prevalent risks and the most efficient mitigation strategies from the collected checklist and present them in the form of a table. The benefits of this research are that the user can easily have a glance at the checklist and identify the risks and mitigation strategies. They can also identify the risks that might occur in the industry and take some preventive measures to minimize them. They can also have a glance at the top five risks and mitigation strategies and can prioritize their choices in terms of which risks has to be attended to first or which mitigation strategy can be used first in the event of a risk.

III. RESEARCH DEFINITION AND PLAN

Research objective(s):

The main goal of this research is to identify the current security issues related to data security in cloud computing and to identify the mitigation strategies which are used to crack these issues. This study also has some sub-goals. They are

- Understanding the prominence of data security aspects in cloud computing
- To get an idea on the current data security issues in cloud computing
- Identifying the mitigation strategies and techniques present.
- To formulate two checklists, one with the risks or challenges and the other with mitigation strategies or techniques.
- Validate the checklists with the help of a survey
- To identify the top five risks those are faced by the industry in the current scenario from the identified checklist.
- To identify the top five mitigation strategies those are used by the industry from the current checklist.

Research questions / hypothese(s):

The following research questions were formulated to continue this research

- R.Q.1: What are the security issues or risks related to data security in cloud computing?
- R.Q.2: What are the mitigation strategies or techniques that are required to crack these risks?
- R.Q.3: Are the risks and mitigation strategies that are identified related to the current security issues in cloud computing?
- R.Q.4: What are the top five risks faced by the industry and the top five mitigation strategies used from the identified list?

Research method:

We have followed the following research method to crack the formulated research questions

- To answer the research questions R.Q.1 and R.Q.2, we have conducted a literature review on the current security issues in cloud computing related to data security and the mitigation strategies in practice in order to crack the issues or challenges. After performing the literature review, we could identify several risks related to data security in cloud computing and several mitigation strategies that are in practice. These risks and mitigation strategies are formulated as two separate checklists.
- To answer the research questions R.Q.3 and R.Q.4, we have chosen to conduct a survey. The survey was conducted among a number of experts and individuals from the industry who have vast experience on the field of cloud computing. A questionnaire was developed in order to perform this survey. This questionnaire was sent to the respondents and their response was recorded.

IV. RESEARCH OPERATION

Checklist:

After conducting the literature review, we could identify 19 risks and 17 mitigation strategies. These risks were identified by various authors with the help of different methods and approaches. After the identification of these risks and mitigation strategies, these were documented as two separate checklists. The first checklist consists of the risks or challenges along with their description. The second checklist consists of the mitigation strategies and their brief description.

The identified risks are concerned with data security in cloud computing. These risks include the risks due to proneness to phishing attacks, risks due to spoofing, risks due to vulnerable storage of data, risks due to elevation of privileges, risks due to interception of data during transmission, risks due to lack of secured APIs, risks due totampering, risks due to inconsistency of data and so on. These risks are documented in a checklist. The description or the context of the risks is provided along with their respective risks.

The identified mitigation strategies or techniques are mentioned to counter the risks that occur in data security aspects in cloud computing. These mitigation strategies include use of security policies, effective encryption, ensuring the security of software, proper authentication and session management, ensuring the security and validity of the firewall certificate, providing security updates and patches for security loopholes and so on. A total of 17 mitigation strategies were identified. A short description of the mitigation strategies was also provided along with their respective mitigation strategies.

Validation of the Checklists:

In order to answer the research question R.Q.3 and R.Q.4, we have chosen to perform a survey. For the survey, we have approached 40 individuals comprising of subject experts and individuals from the industry. Each of them has an experience of 3-12 years in the area of cloud computing specifically confined to data security. The reason for choosing a survey is that it is well suited method for our research. The basic idea of our research is to validate the checklists and this can also be performed through an experiment or an interview. Since, we need to collect more respondents to validate the results and interview cannot be performed in our time constraint, and experiment is confined to a single scenario, we chose to perform a survey. In order to perform this survey, we have followed a two-step approach. Step 1 is intended to crack the research question R.Q.3 and Step 2 is intended to crack the research question R.Q.4. The steps of the survey in parts are described in the following sections.

Step 1:

In the first part, a questionnaire was distributed among the respondents. Based on the questionnaire, the respondents were asked to select the risks or mitigation strategies from among the checklist. The questionnaire is as follows

1. Which of the risks specified in Checklist-A are related to data security in cloud computing?
2. Which of the risks specified in Checklist-A specify the current problems related to Data security in cloud computing?
3. Which of the mitigation strategies specified in Checklist-B are currently in practice to minimize the risks related to data security in cloud computing?

The respondents were asked to select the options which they felt were the best for the given questions from the checklist. The response is recorded and processed in the following manner.

For example, if a specific respondent had voted for 'x' number of options for a particular question. The percentage validity 'P' of the checklist is calculated by using the formula

P= (x/19)*100 [for checklist –A] P = (x/17)*100 [for checklist- B]

Based on the value of P, we designed a range of values for the options we have chosen. The range and the options are given in the following table

OPTION	RANGE
A : Very Much Related	80 ≤ P ≤ 100
B : Fairly Related	60 ≤ P < 80
C : Slightly Related	30 ≤ P <60
D : Poorly Related	0 ≤ P < 30

Table: Options and Range

So using the mentioned formula and the table, we have categorized the individual options for each respondent. Then, the option, as mentioned in the table 3 is calculated for each respondent. Collectively, the options and the percentage of respondents who have voted for the options were graphically represented.

Step 2:

In order to crack the research question R.Q.4, a survey was conducted among the respondents. The respondents were provided with the risks and mitigation strategies that were obtained from the checklist formed. The respondents were asked to vote on the risk which they felt were the most occurred problem in the industry. Similarly, they were also asked to vote on the mitigation strategy which is prevalently used in the present industry to effectively crack their problems. The number of votes received by each risk and mitigation technique was counted. The results were graphically presented.

With the help of results obtained, we could identify the top five risks related to data security in cloud computing that are prevalent in the industry. We could also identify the top five mitigation strategies or techniques that are implemented by the industries to counter the risks or challenges that they face.

V. DATA ANALYSIS AND INTERPRETATION

In order to crack the research questions R.Q.1 and R.Q.2, we have conducted a systematic literature review. This method helped us in the identification of literature which enabled us to identify the risks that are prevalent in cloud computing related to data security. The risks were collected from various papers and from various authors. These risks were documented as a checklist. A short description of the risks is also provided alongside each risk. The checklist A is presented in the appendix.

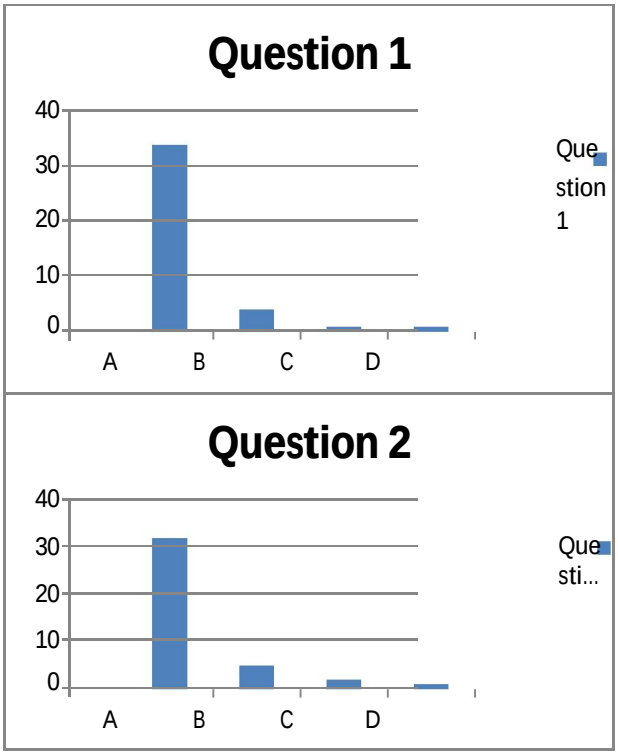
We also could identify a set of mitigation strategies that are used in order to counter the risks related to data security in cloud computing. The mitigation strategies are also presented as a checklist B. A short description of the mitigation strategies was also provided alongside each mitigation strategy. The checklist-B is presented in the appendix.

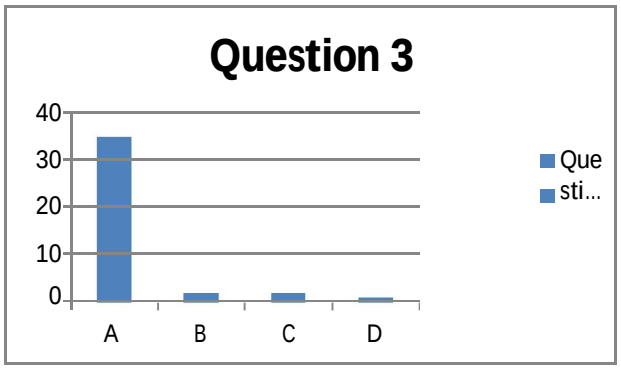
In order to crack the research questions R.Q.3 and R.Q.4, we have conducted a survey. . For the survey, we have approached 40 individuals comprising of subject experts and individuals from

the industry. Each of them has an experience of 3-12 years in the area of cloud computing specifically confined to data security. This survey was conducted in two steps. The details or the method of survey was explained in the previous section. We now present the results of the survey.

The result of step 1 of the survey is presented below.

Options	Value
A	Very much related
B	Fairly related
C	Slightly related
D	Poorly related





On analyzing the results graphically, we could observe that almost all of the respondents have responded that the risks and mitigation strategies collected in the checklist are very much related to the present risks in the industry and the mitigation strategies practiced in the industry. Very few of them have opined that the risks and mitigation strategies in the checklist have a slight and fair relation to those in the industry. Only one or two of them opined that they have a poor relation to those in the industry. Therefore, with the help of this survey, we could confirm that the risks collected in the checklist are the present trending problems in the industry related to data security in cloud computing and the mitigation strategies collected in the checklist are the present methods used in the industry to counter the problems.

The result of step 2 of the survey is presented below. Due to space constraint, only the top five polls are presented.

Top Five Risks

A	Risk due to vulnerable storage of data
B	Risk due to the loss of control on the data present in the cloud
C	Risk due to the access by unauthorized users
D	Risk due to the lack of proper authentication and authorization to the data
E	Risk due to lack of secured APIs and Interfaces

Top Five Mitigation Strategies

A	Effective encryption
B	Providing security to the storage of data
C	Implementing efficient plans for security
D	Use of security policies
E	Proper authentication and session management

On analyzing the results graphically, we could identify the top five risks and mitigation strategies from the checklist which are the present trending risks and present efficient mitigation strategies in the industry related to data security in cloud computing. This was done based on the number of votes received by the risks and mitigation strategies polled in by the respondents. After identifying the top five risks and mitigation strategies, they were tabulated as follows.

VI. DISCUSSION

We now present our discussion and analysis on the research we have performed. We start with presenting our research method in short. With the help of a literature review, we could identify the present problems in the industry related to data security in cloud computing. We also could identify the mitigations strategies used to counter the risks. We have presented them in two checklists: checklist-A for risks and checklist-B for mitigation strategies. In order to validate the checklists, we have conducted a survey. With the help of the survey, we could validate that the risks and mitigation strategies that were collected in the checklist are the present trending risks that are faced by the industries and the mitigation strategies used by the industries in the current scenario. We also have conducted a survey with the same respondents to identify the top five risks and mitigation strategies from the checklist. After analyzing the results of the survey, we have presented the top five risks and mitigation strategies from the checklist in a separate table.

Contributions:

- We could identify the risks that are related to data security in cloud computing and validated them. The novelties of result in comparison to previous research are that we have presented them in a checklist which is an effective way of presenting the results. A viewer can easily have a glance at the checklist and easily understand its contents. We have validated them to those risks in the industry. Therefore, industries can make use of this checklist and get an idea on the main risks that are prevalent in data security and can take preventive measures.

- We could identify the mitigation strategies that are related to data security in cloud computing. The novelties of this result in comparison to previous research are that we have presented them in a checklist along with the risks, which is an effective method of presenting the results. A viewer can have a glance at the checklists and easily identify and understand the mitigation strategies. We have validated them to those mitigation strategies practiced in the industry. Therefore, industries can make use of this checklist and can use the mitigation strategy that is most suitable for them in terms of the usage or the problem.

- We could identify the top five risks and mitigation strategies from the checklist and presented them in a table. The novelties of this result in comparison to previous research are, we could see that

there wasn't much effort done to identify the most prevalent, common or top risks that affect the industry and the mitigation strategies that can be used extensively to mitigate the common risks. With the help of this table, the industries can have an idea on the top five risks and mitigations strategies so that, they can be prepared or take steps in order to prevent the top risks which affect the industry mostly. They can also identify and make use of the top five mitigation strategies which are widely used and which are very affective in mitigating the risks in data security.

Threats to validity

There are various threats that may occur during this research. They are described as follows

- Internal Validity is the *"The extent to which the design and conduct of the study are likely to prevent systematic error"[19]*. One of the threats is lack of enthusiasm and interest in the respondents. This can drastically bring down the efficiency of the survey. To avoid this threat, we have selected the respondents who were fully motivated and fully interested to participate in the survey.

- We have considered the risk of reliability of the survey. This might arise due to the problem of the respondents in understanding the language of the questionnaire or problem in understanding the questions or the design of the questionnaire. To avoid this, we have presented short questions in an understandable and simple way. We only have presented the respondents with three to four questions so that they might not face a problem in understanding and analysing the questionnaire.

- We have considered the risk that the documented result may not be applicable to all the industries. In order to avoid this risk, we have specified that this research is confined to data security issues and mitigation strategies. Therefore, the industries facing the problems with data security can make use of these checklists.

- We have considered the validity of the population as a threat. This is a problem when the selected population size for the survey is very small for generalizing the result. To avoid this, we have considered a very good number of people for the survey in order to generalize the results.

VII. SUMMARY AND CONCLUSIONS

We have tried to identify the risks and mitigation strategies related to data security in cloud computing. These were presented as two separate checklists. The validation of the checklists was done with the help of a survey and the top five risks and mitigation strategies were tabulated. Since cloud computing is an area of vast research and still there is a much broader scope for improvement, as a future work, we would recommend to collect multiple mitigation strategies for each risk and then prioritize them. Also, more mitigation strategies and frameworks can be tested in cloud environment.

REFERENCES

Almulla, S.A.; Chan YeobYeun; , "Cloud computing security management," *Engineering Systems Management and Its Applications (ICESMA), 2010 IEEE Second International Conference*, pp.1-7, March 30-April 1 2010.

B. A. Kitchen ham, S. Charters, “Procedures for Performing Systematic Literature Reviews in Software Engineering”, EBSE Technical Report, Software Engineering Group, School of Computer Science and Mathematics, Keele University, UK and Department of Computer Science, University of Durham, UK, 2007.

Blumenthal, M.S.; , "Hide and Seek in the Cloud," *Security & Privacy, IEEE* , Vol. 8, No. 2,pp.57-58, March-April 2010.

Chang-Lung Tsai; Uei-Chin Lin; Chang, A.Y.; Chun-Jung Chen;, "Information security issue of enterprises adopting the application of cloud computing," **etworked Computing and Advanced Information Management (*CM), 2010 IEEE Sixth International Conference*, pp.645-649, Aug. 2010.

Chow, R.; Golle, P.; Jakobsson, M.; Shi, E.; Staddon, R.; Molina, J.;, "Controlling data in the cloud: Outsourcing computation without outsourcing control." Proceedings of the 2009 ACM workshop on Cloud computing security, pp.85-90, 2009.

Clarke, Roger; , "User Requirements for Cloud Computing Architecture," Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference, pp.625- 630, May 2010.

CSA, "Top Threats to Cloud Computing V1.0," http://www.cloudsecurityalliance.org/topthreats/csathreats.v 1.0.pdf,

de Chaves, S.A.; Westphall, C.B.; Lamin, F.R.; , "SLA Perspective in Security Management for Cloud Computing," **etworking and Services (IC*S),2010 Sixth International* 43 *Conference*, pp.212-217, March 2010.

Delettre, C.; Boudaoud, K.; Riveill, M.; , "Cloud computing, security and data concealment," Computers and Communications (ISCC), 2011 IEEE Symposium on , vol., no., pp.424-431, June 28 2011-July 1 2011 doi: 10.1109/ISCC.2011.5983874

Gruschka, N.; Iacono, L.L.; , "Vulnerable Cloud: SOAP Message Security Validation Revisited," Web Services, 2009. ICWS 2009. IEEE International Conference, pp.625-631, July 2009.

Itani, W.; Kayssi, A.; Chehab, A.; , "Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures," Dependable, Autonomic and Secure Computing, 2009. DASC '09. Eighth IEEE International Conference on , vol., no., pp.711-716, 12-14 Dec. 2009 doi: 10.1109/DASC.2009.139

Kai Hwang; Kulkareni, S.; Yue Hu;, "Cloud Security with Virtualized Defense and Reputation-Based Trust Management," Dependable, Autonomic and Secure Computing, 2009. DASC '09. Eighth IEEE International Conference, pp.717-722, Dec. 2009.

Kienle, H.M.; Muller, H.A.;, "Research challenges in management and compliance of policies on the web," *Web Site Evolution, 2008. WSE 2008.10th International Symposium, IEEE*, pp.83-92, Oct. 2008.

Komashinskiy, D.; Kotenko, I.;, "Malware Detection by Data Mining Techniques Based on Positionally Dependent Features," *Parallel, Distributed and*etwork-Based Processing (PDP), 2010 18th Euromicro International Conference*,pp.617-623, Feb. 2010.

Li, H.; Dai, Y.; Tian, L.; Yang, H., "Identity-based Authentication for Cloud Computing," Cloud Computing. Proceedings First International Conference, CloudCom. pp.157-66. 2009.

Martignoni, L.; Paleari, R.; Bruschi, D.;, "A Framework for Behavior-Based Malware Analysis in the Cloud," Information Systems Security. Proceedings 5th International Conference, ICISS, pp.178-92, 2009.

Mather, T.; Kumaraswamy, S.; Latif, S.;, Cloud Security and Privacy An Enterprise Perspective on Risks and Compliance: O'Reilly, 2009

Minqi Zhou; Rong Zhang; Wei Xie; WeiningQian; Aoying Zhou; , "Security and Privacy in Cloud Computing: A Survey," *Semantics Knowledge andGrid (SKG), 2010 Sixth International Conference*, pp.105-112, Nov. 2010.

Owens, D.;, "Securing elasticity in the cloud: Elastic computing has great potential, but many security challenges remain," Queue, Vol. 8, No. 5, pp.1-7, 2010.

Paquette, S.; Jaeger, P. T.; Wilson, S. C.;, "Identifying the security risks associated with governmental use of cloud computing," Government Information Quarterly, Vol. 27, No. 3, pp. 245-253, 2010.

Sabahi, F.; , "Cloud computing security threats and responses," Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on , vol., no., pp.245-249, 27-29 May 2011

Saripalli, P.; Walters, B.; , "QUIRC: A Quantitative Impact and Risk Assessment Framework for Cloud Security," Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference, pp.280-288, July 2010.

Wang, C.; Wang, Q.; Ren, K.; Lou, W.; , "Towards Secure and Dependable Storage Services in Cloud Computing," Services Computing, IEEE Transactions on , vol.PP, no.99, pp.1, 0doi: 10.1109/TSC.2011.