Security Analysis in Context-Aware Distributed Storage and Query Processing in Hybrid Cloud Framework

Geremew Begna
Department of Electrical Engineering and Computer Science
Howard University, Washington, DC, USA
gbegna@gmail.com

Danda B. Rawat
Department of Electrical Engineering and Computer Science
Howard University, Washington, DC, USA
db.rawat@ieee.org

***Abstract*-Recent studies have shown that several government and business organizations experience huge data breaches. Data breaches increase in a daily basis. The main target for attackers is organization sensitive data which includes personal identifiable information (PII) such as social security number (SSN), date of birth (DOB) and credit card /debit card (CCDC). The other target is encryption/decryption keys or passwords to get access to the sensitive data. The cloud computing is emerging as a solution to store, transfer and process the data in a distributed location over the Internet. Big data and internet of things (IoT) increased the possibility of sensitive data exposure. Most methods used for the attack are hacking, unauthorized access, insider theft and false data injection on the move. Most of the attacks happen during three different states of data life cycle such as data-at-rest, data-in-use, and data-in-transit. Hence, protecting sensitive data at all states particularly when data is moving to cloud computing environment needs special attention. The main purpose of this research is to analyze risks caused by data breaches, personal and organizational weaknesses to protect sensitive data and privacy. The paper discusses methods such as data classification and data encryption at different states to protect personal and organizational sensitive data. The paper also presents mathematical analysis by leveraging the concept of birthday paradox to demonstrate the encryption key attack. The analysis result shows that the use of same keys to encrypt sensitive data at different data states make the sensitive data less secure than using different keys. Our results show that to improve the security of sensitive data and to reduce the data breaches, different keys should be used in different states of the data life cycle.**

Keywords- cloud computing, data breaches, sensitive data, data security, encryption, encryption key, data state, data classification

I. INTRODUCTION:

Today, more organizations are moving their Information system to the cloud computing environment to store, process, retrieve, and share data with more people in different locations. As defined by [1], cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources(e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort. The cloud model is composed of five essential characteristics, three service models, and four cloud deployment models. The amount of data producing is growing exponentially due to technological advances. For example, the increasing number of people, devices, and sensors that are now connected by digital networks is called Internet of Things (IoT) is the main reason for the vast amount of data we generate every day [3, 4]. The large data is brought new term called big data. This brought more security challenges as most of organizations worries to store the large volumes of data including personal data and organizational strategic information in cloud, processes the stored sensitive data over the internet, and retrieves sensitive data from stored location (Cloud).

Sensitive Data/ Personal Data

Sensitive data is a term that includes identifiable data, but also extends to information that may be considered private, or to have societal and economic impact on sequences if improperly disclosed [2, 3]. Sensitive data could include government assigned identification, financial data, medical data and history, employee records, organization business strategic information, military information etc. The misuse of personal identifiable information (PII) contributes to a range of losses that affect how an individual can participate or benefit from financial, employment, social, medical, and other activities. For example, stolen PII can be used for criminal activities and to misdirect law enforcement investigations to wrong person.

Data Breaches Problem

Data breaches is an incident in which sensitive, protected or confidential information is released, viewed, stolen or used by an individual who is not authorized to do so. In recent years, both government and business organizations have experiencing large data breaches that have made the PII of millions of individuals available to those who would then use the information to commit identity fraud. As technology advance, data breaches increase. According to the 2017 data breach year-end review released by the Identity Theft Resource Center (ITRC) [5] and Identity force [6], data breaches are increasing from year to year. For example, the number of U.S data breach incidents hit a new record high of 1,579 breaches, a drastic rise of 44.7 percent increase over the record for 2016.

Industry Sectors Affected by Data Breaches

ITRC data breach report [5] also identified top five industry sectors that are highly affected by data breaches. These are business, medical/healthcare, banking/financial, education, and government /military. Among the five industry sectors, the business category takes the first rank in total number of breaches followed by medical /healthcare industry in the second place. The banking/credit/financial sector takes third place. The remaining two sectors, educational and government/military are in fourth and fifth places respectively. Fig.1 below shows the data breaches by industry sector for the year 2015-2017.

978-1-7281-0554-3/19/$31.00©2019 IEEE

	Year-2015		Year-2016		Year- 2017	
Industry Sector	#Breaches	# Records	#Breaches	# Records	#Breaches	# Records
Business	312	16,191,017	495	5,669,711	870	163,449,242
Medical/Healthcare	276	121,629,812	376	15,942,053	374	5,062,031
Banking/Credit/Financial	71	5,063,044	52	72,262	134	3,122,090
Government/Military	63	34,222,763	72	13,869,571	74	5,903,448
Educational	58	759,600	98	1,048,342	127	1,418,258
Totals for All sectors	780	177866236	1093	36601939	1579	178955069

Fig.1.Breaches by industry sectors for year-2015-2017

Methods used for Data Breaches

Attacker use different methods to access sensitive data. According to the ITRC report [5], the seven major methods used for the attacks are hacking, unauthorized access, insider theft, data on the move, accidental exposure, employee error, and physical theft. The Fig.2 below summarizes the data breaches by type.

Type of Attack	#Breaches	#Record
Hacking	940	167,549,245
Unauthorized Access	170	467,215
Insider Theft	83	959,909
Data On The Move	34	105,504
Accidental Exposure	101	8,018,791
Employee Error	164	145,927,550
Physical Theft	71	1,399,925

Fig.2. Data breaches by type of attack.

Sensitive Data Affected During Data Breaches

From the ITRC report ,we have identified the major information compromised during data breaches . According to the report, the affected PII are social security number (SSN), credit/debit card number (CCDC), protected health information (PHI),driver's license ,financial accounts, email/password/user name other/undefined type of records. Among the sensitive data exposed, SSN and CCDC are the two most exposed PII.

Data State and Vulnerabilities

Sensitive data can be compromised at any state during data life cycle. [7, 13] classified the state of data in its process life cycle into three states. These are data-at-rest, data-in-transit, and data-in-use. These classifications reflect the states of data from the time it is created or stored, processed through computational resources, such as in CPU components and main memory, and transmitted over the internet to the cloud.

In recent years, data leakage attacks have been achieved on sensitive data at all these three data states: First, it is important to understand the three data states and security concerns.

Data-At-Rest-In this state, data is stored in storage devices. For example, the internal hard disk of a computer is permanent storage for data. It is a representation of data when it is stored in storage devices. In cloud computing, there is a concern about the location transparency of data due to the complexity of the cloud infrastructure. Because cloud consumers do not exactly know where their data is stored in cloud systems. This situation causes uncertainty among organizations considering the use of cloud systems, worrying about who deals with sensitive data. Moreover, customers are also worried about malicious insiders who co-locate on the same storage. Recent Studies show data attack when data is at rest including in the cloud storage. For instance, [8] identified a hard disk-based covert channel attack against arbitrary virtual machines (VMs) such as Amazon EC2.

Data-In-Transit-This state describes the transfer of data over networks, which is delivered from a source to a destination. Data will be at more risk in this media due to spying attack, which is able to inspect packets. [9] showed a possible leakage attack against network packets; mainly, when they are carrying credential information and moving through SSH. To tack data attack at this state, the authors developed a hidden markov model (HMM) and key sequence prediction algorithm to guess correct pressed keys, which forms the password by victim.

In cloud systems, software-as-a-service (SaaS) enables web technology to deliver applications as a service to end users. Data in use state also represents request/response between client and server while they communicate. The sensitive data needs to be protected by encryption from online spying attack. However, this does not stop attackers from malicious intentions. Attackers can still analyze pocket size and timing to leak sensitive information.

Data-In-Use-In this state, data is loaded into computational resources (mainly in to the CPU components such as registers, caches and main memory). When data is loaded from storage devices to CPU registers, it must pass through different hierarchical buffers, including main memory and CPU caches prior to reaching the registers, where it is prepared for operations such as read and write. According to the applications, data might have different alignment and organizations while loading into main memory regarding the data types. These processes expose sensitive data to the attacks. In a multitasking environment, physical resources are shared between processes indifferent layers, such as OS (page sharing) and application (shared libraries), the data frequently being used by those resources. Data security in cloud computing is using those resources in the case that a region of memory is shared between processes. Recent work has shown vulnerabilities of data-in-use in multitasking systems in both OS sharing [10], the achieved memory deduplication attack and application (shared library). Attacks on data at all three data states is also discussed briefly in [11].

From the above data states review, we know that data breaches can happen during any states of data life cycle, Hence,

protecting sensitive data at all states (i.e. storage, transit and retrieval or use) is very important. The purpose of this research is to examine previous sensitive data protection approaches and propose a better approach that helps individuals, businesses and government organizations to protect their sensitive data.

II. Related Work

There are several related works in cloud computing that we have reviewed in support of this research. Among the related works [12] discusses issues related to data security and privacy aspects in cloud computing, such as data integrity, data intrusion, service availability. The authors proposed a multi-clouds database model (MCDB) which is based on multi-clouds service providers instead of using single cloud service provider. In addition, the paper discusses the architecture of the proposed MCDB model and describe its components and layers.

Another paper [13] presents about the personal data usage in cloud and the several security concerns such as lack of user control, the non-compliance with the user's preferences and regulations, the difficulty of the data flow tracking, etc. In particular, the paper discusses the unsolved problem to ensure customers data usage policies regardless of who accesses the data, how data is processed, where the data is stored, transferred and duplicated. The paper calls for two requirements to be satisfied. First, data should be handled in accordance with both owners' preferences and regulations policies whenever it exists in the cloud and throughout its lifetime. Second, a consistent data flow tracking should be maintained to follow up the data derivation. To address these issues, the researchers proposed a hybrid approach to protect private data in the cloud. The paper proposed the PriArmor data content for self-defending data when it data stored or transferred in the cloud .According to the paper , the PriArmor agent acts as an armor for privacy protection when private data is processed by the cloud-based services.

Another related paper [14] discusses the lack of traditional security mechanisms such as firewalls, virtual private networks (VPNs), and intrusion detection systems/intrusion prevention systems (IDSs/IPSs) to prevent the leakage of sensitive data. The paper presents a model called data leakage prevention systems (DLPSs) to overcome the deficiency of traditional tools in protecting sensitive data. The researcher proposed hybrid symmetric-asymmetric encryption to prevent against data leakage. The proposed encryption method ensures that all confidential or sensitive documents of an organization should be encrypted so that only users with access to the decrypting keys can have access. The researcher classified data into confidential and non-confidential using the method called Naïve Bayes Classifier

This research is an effort to improve previous works. particularly the work of [14] to secure sensitive data by preventing the leakage of the sensitive data. The solution we propose is a better approach to protect the disclosure of sensitive data problems at all data states, data-at-rest, data-in-transit, and data-in-use. This paper is also a continuation of authors' previous work [15] to investigate how sensitive data can be protected in a distributed and hybrid cloud computing framework.

III. System Model

The secure model we propose is a better approach to control personal data and protect privacy. The model works in such a way that organizations data should be classified initially based on its context. Data that contain highly personal information or PII is classified as high sensitivity (HS), data that contain medium sensitive information is classified as medium sensitivity (MS), and data that contain information for general public as low sensitivity (LS). Based on its sensitivity level, access to data

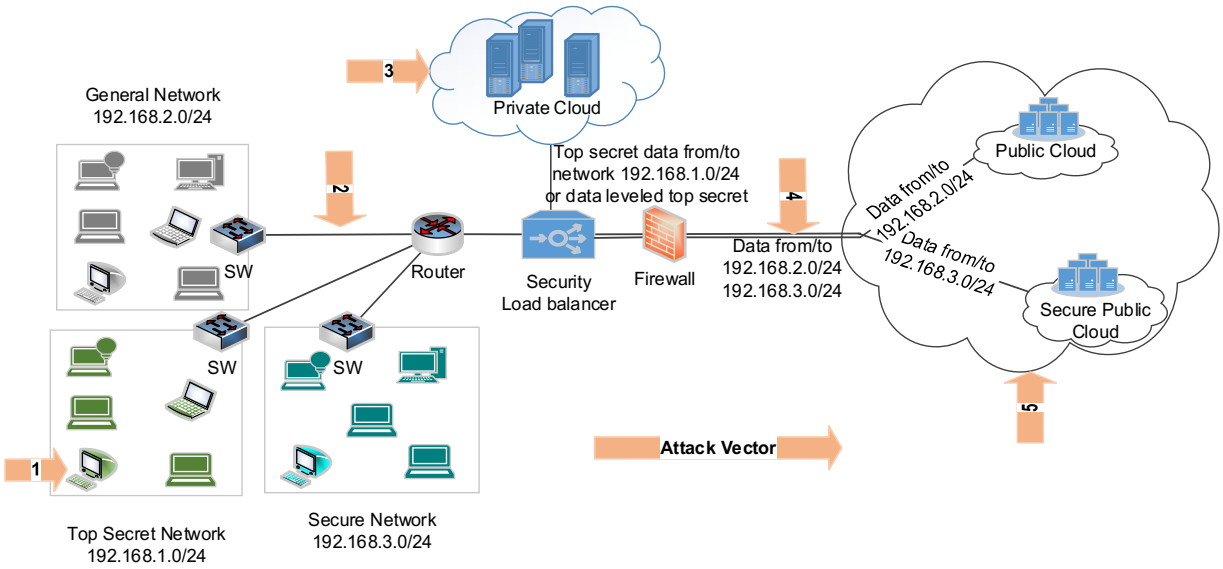

Fig. 3. Typical System Model with Internal, Private and Public Cloud Scenario

should also be determined. In other words, who should get access or who should not get access. Normally, data that contain PII or considered as private should be encrypted at the beginning stage when entered and stored in the local network storages. Unlike sensitive data, data for general public or the non-sensitive may not require to be encrypted. During data deployment, the model algorithm (security Load balancer) determines the destination of the data in the cloud system. The security load balancer checks the sensitivity level provided to the source data during data classification and sends the data to either the private cloud or public (public, secure public) cloud models. The algorithm of the system model is discussed under section E below.

To demonstrate how to protect sensitive data in business or government organizations, we represent typical business organization's local network and cloud IT infrastructure. The organization has different departments, local networks, and employees working in different departments. Employees working in different departments process data of different sensitivity levels. The organization has different networks (top secure, secure, and general). The organization also needs different user roles to assign them access privileges. The access privileges should be granted to system users based on their work responsibilities.

To better manage the big data and take advantage of the cloud, the organization needs to deploy its IT infrastructure to cloud services. The main requirement is to store sensitive data such as employees and customers SSN, DOB, salary information and customer credit card information to store stored the sensitive data from top secure network in the private cloud, medium secure data from the secure network in the secure public and data for the public from general network in the public cloud's environment respectively. Data stored at these different locations can also be transferred from one node (location) to other nodes (locations) and retrieved for further processes and uses.

One of the current challenges of business organization is how to keep sensitive data from unauthorized users or attackers while data is at all states (at rest, in transit, and in use) and how to prevent data breaches? The model we propose presents better solution. This paper discuss how data should be categorized at its original stage, how to encrypt, what keys should be used to encrypt data at different states ,what data to be encrypted and which cloud model (private or secure public or public) should be used to store data base on data sensitivity level and source .

A.Context-Aware Data Classification

Our main goal in this study when considering data classification is to identify and specify the sensitivity of organization data. For example, is the data freely and publicly available, or for internal use only, or is it sensitive information that only certain personnel should be allowed to access the data.

Data classification is the first step in our data security approach. The classification is based on data sensitivity level, organization data retention policy and standards, and severity level. We determined to take the following three classifications approach discussed in [2] with basic class definitions:

High Sensitivity (HS)-Information that would cause major damage to the organization or subject if disclosed externally from the organization. The damage referenced in this description includes exposure to lawsuit, compromise to security of assets, reputation of organization and its associates, elimination of competitive advantage, and violation of regulation, industry standard or corporate policy, sensitive or identifiable through regulation, industry standard or corporate policy. PII such as SSN, DOB, CCDC are under this data classification category.

Medium Sensitivity (MS)-Information that would cause minor damage to the organization or subject if disclosed externally from the organization. Data that is for internal use such as regulation, industry standard or corporate policy can be grouped under this category.

Low Sensitivity (LS)-Information that is publicly available through other civil sources, or that has been specifically designated as public information by internal regulation or corporate policy. for example, information on company's website.

B. Classifications of Networks

Depending on the context or sensitivity level of the data, organizations should also determine their local network or server types, where data should be stored and processed. For example, data with high sensitivity level should be stored in top secure network, data with medium sensitivity in secure network, and data with low sensitivity level in general network. Therefore, we can classify local networks as top secure network (TN), secure network (SN), and general network (GN).

C. Selection of Cloud Models

We propose a hybrid cloud computing framework. Because the hybrid cloud model comprises of private and public clouds so that we can store data that is highly sensitive in private cloud ,data that is medium sensitive in secure public cloud and data for general public in the secure public cloud respectively. The three cloud models and purposes are discussed as follow:

Public Cloud-is used to store information that is publicly available for use by General public and from the Local general network

Secure Public-is used to store medium sensitive information for internal use only and sensitive data from the secure network.

Private Cloud-is used to store the highly sensitive data class which includes information that is designated explicitly as sensitive or identifiable through regulation, industry standard

D. Users Role and Permission

Before providing permission, users need to be assigned to one or more roles designated by the organization. The assignment is based on users roles , responsibilities and data the users must access. In general terms, identify users' roles and data each user or group could access with assigned role or roles within the organization local network system. Then, elevate same permissions and roles when the user get access to cloud .

Permissions are used to define who can access specific objects or data within the organization networks and in the cloud. Permissions can also define how a user interacts with the object or data. Access control and how to assign permission to different roles is briefly discussed in the author's thesis work [16].

E. The Algorithm

The following short algorithm or steps demonstrates how our system model works

1. *START*
2. *Identify data sensitivity level, HS, MS or LS*
3. *Identify the source network, TN, SN, GN*
4. *If data is HS, Then, encrypt and store in TN*
5. *ELSE IF data is MS, then, encrypt and store in SN*
6. *ELSE store the data in GN*
7. *IF data is HS and source network type is TN, Then, encrypt and send the data to Private cloud*
8. *ELSEIF data is MS AND source network is SN, Then, encrypt and send the data to Secure public cloud*
9. *ELSE send the data to Public cloud*
10. *END IF*

IV. Securing Sensitive Data in Hybrid Cloud Computing

To prevent the attack of sensitive data in different data states, we propose a model discussed under section III. In addition to the host-based firewall and host-based intrusion prevention/defection systems at the end device, we propose data sensitive data encryption to defend any attacks of sensitive data at Local level. When the data is generated at the end device, it is labeled with sensitivity levels before encryption. The encryption key used to encrypt data between the local network and security load balancer is generated by using the Diffie-Hellman algorithm [15].

Securing Data-At-Rest-Encrypting hard drives is also one of the best ways to ensure the security of data at rest. Other steps can also help is storing individual data elements in separate locations such as private cloud to decrease the likelihood of attackers gaining enough information to commit fraud or other crimes. Creating strong password is the best practice to protect data from unauthorized access. We can also use different encryption key to encrypt different data stored in different networks.

Securing Data-In-Transit-The data being transmitted between the end device and security load balancer is encrypted with Diffie-Hellman key which helps prevent the sensitive data attack while data is in motion. Similarly, data is encrypted before it leaves for private cloud or secure public cloud storages using another Diffie-Hellman key [17] or different key generated by other algorithm. This approach prevents sensitive data from another attack in the middle. The best way to ensure that sensitive data remain confidential is to transmit the data through an encryption platform that integrates with the existing systems and workflows.

Moreover, when the data is requested (query retrieval) from where it is stored in the cloud (secure, public or private), the query statement is converted into its encrypted version. The security load balancer converts the received query into its encrypted query statement and communicates to the cloud on behalf of the end user. These steps also help to secure data in motion at different stages. Note , the paper discusses the impact of using the different and same encryption keys and which one is better to secure sensitive data under the next section V.

Securing Data-In-Use-For the integrity of sensitive data in use, it must be accessible to those who need it. Of course, the more people and devices that have access to the data, the greater the risk that it will end up in the wrong hands at some point. Therefore, the better ways to securing data in use are to control access as tightly as possible and to incorporate some type of authentication to ensure that users aren't hiding behind stolen identities. Organizations should also track and report relevant information so that they can detect suspicious activity, diagnose potential threats, and proactively improve security. For example, disable accounts if the account failed after number of failed logins attempts or send a warning sign that a system is under attack to the system administration.

V. Mathematical Analysis

We are going to figuring out the probability that the legitimate user and the attacker have the same encryption key. To solve this problem, we used the birthday paradox scenario (18). For this analysis, we also apply rule of independent event probability [19].We use the following variables: N represents the number of generated unique keys, n represents the number of known encryption keys to the attackers, P(S) = probability that the key is known to the attacker, P(D) = probability that the key is not known to the attacker. The probability that n keys used are to be the same can be calculated as the permutation.

P (no n same keys) = $P(D) = N!/((N-n) * N^n)$ (1)

and the probability that n keys will be the same is

P (n same keys) =$P(S) = 1 - N!/((N-n) * N^n)$ (2)

For instance, when a key is known to the attacker, then, n = 2.

P1 (key is known to attacker) =$1 - N!/((N-2) * N^2)$ (3)

and the probability that n keys will be the same as P(S)

P2(key is NOT known to attacker) =$N!/((N-2) * N^2)$ (4)

Using the above equations, we can calculate the probability that encryption key is known to the attacker and not known to the attacker at local network, private cloud ,and secure public cloud. For example, for a local data encryption, for a given number of keys Ni, the probability of a key used to encrypt data is known to the attacker can be calculated as follows

$$P(Si) = 1 - Ni!/((Ni-2) * Ni^2) \quad (5)$$

and the probability of a key used to encrypt data at local network is NOT known to the attacker is

$$P(Di) = Ni!/((Ni-2) * Ni^2) \quad (6)$$

Similarly, for private cloud communication encryption, for a given number of keys Npc, the probability of a key used to encrypt sensitive data in the private cloud is known to attacker is

$$P(Spc) = 1 - Npc!/((Npc-2) * Npc^2) \quad (7)$$

and the probability of a key used to encrypt sensitive data in private cloud is NOT known to the attacker is

$$P(Dpc) = Npc!/((Npc-2) * Npc^2) \quad (8)$$

Finally, for secure public cloud communication, the probability of a key used to encrypt data is known to attacker is

$$P(Ssc) = 1 - Nsc!/((Nsc-2) * Nsc^2) \quad (9)$$

where Nsc is the number of keys used in secure public cloud. The probability of a key used is NOT known to the attacker is

$$P(Dsc) = Nsc!/((Nsc-2) * Nsc^2) \quad (10)$$

For public cloud communication, we may not need to encrypt data because the data stored in public cloud should not contain sensitive data and is available for general public.

We can also calculate the probability that sensitive data being hacked or leaked when data is travelling from local network to the private. To get the result , we multiply the probability of a key used to encrypt data is known to the attacker at local network P(Si) and the probability of a key used to encrypt is known to attacker during data is travelling to the private cloud P(Spc)

$$P(SiSpc) = P(Si) * P(Spc) \quad (11)$$

Similarly, the probability that sensitive data being hacked, or the probability of informing being leaked when data is travelling from local network to the secure public cloud is

$$P(SiSsc) = P(Si) * P(Ssc) \quad (12)$$

VI. PERFORMANCE EVALUATION AND RESULT

To evaluate the effectiveness of using the different and same encryption keys to encrypt sensitive data at different data states, we performed mathematical analysis. For the analysis, we used a sample data of 1000 for Ni (total number of generated keys to encrypt data at local network) , Npc (total number of generated keys to encrypt data traveling to private cloud), and Nsc (total number of generated keys to encrypt data moving to secure public cloud). By substituting 1000 for Ni, Npc, and Nsc, we can calculate P(S) = P(encryption key is known to the attacker) and P(D) = P(encryption key is NOT known to the attacker). For example, when a key is known to the attacker, the number n=2 ,then when we substitute 1000 for Ni ,the result is following

$$P(Si) = 1 - Ni!/((Ni-2) * Ni^2)$$
$$P(Si) = 1 - 1000!/((1000-2) * 1000^2) = 0.001 \quad (13)$$

and the probability that key is not known to the attacker

$$P(Di) = Ni!/((Ni-2) * Ni^2)$$
$$P(Di) = 1000!/((1000-2) * 1000^2) = 0.999 \quad (14)$$

We prove that the sum of two events ,the key is known to the attacker P(Si) and key is not known to the attacker P(Di) =1

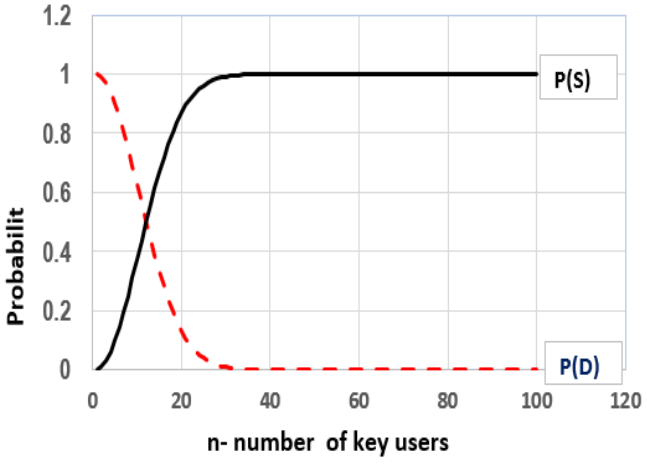

Fig.4. Probabilities of at least one attacker knowing the encryption key of a legitimate users P(S) and not knowing the encryption keys P(D)

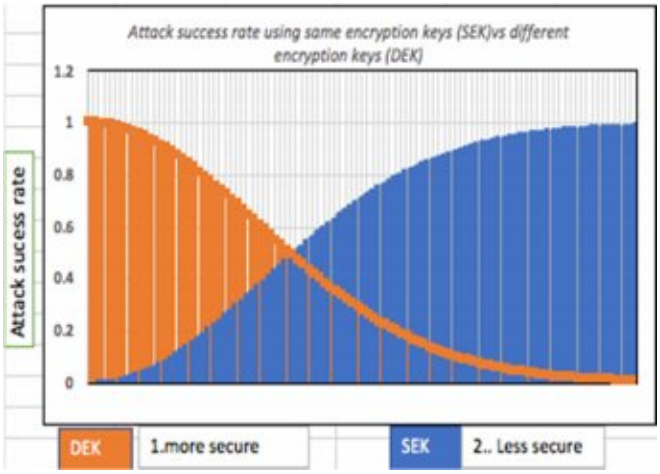

Fig-5.Attack success rate using same encryption keys (SEK)vs different encryption keys (DEK)

Encryption of sensitive data using same encryption key (SEK) for different parts (internal, private cloud, and secure public cloud) in data communication for our system model takes less time when we compare encryption using different encryption key (DEK). The sensitive data that is encrypted using SEK is less secure than using DEK. In other words, the probability of an attacker to break the encryption key is higher when SEK is used to encrypt and decrypt the sensitive data at different stages in our system model. The Fig. 5 above illustrates the attack success rate increases using SEK when compared to that of us-ing DEK.

VII. CONCLUSION AND FUTURE WORKS

The paper has discussed the growing concerns of data breaches in both business and government sectors and how to protect sensitive data or PII at the local network as well as in the cloud environment. Using our proposed model, we illus-trated how data classification, data encryption helps to secure sensitive data when it is at rest, during transferring data from local network to cloud storage. The paper also presents how sensitive data should be secured through its life cycle from local network storage to its destination in the cloud (private or public) using same and different encryption key. Our study result shows that it is better use different keys to secure sensitive data at different states of its life cycle to reduce data breaches.

Our future work will focus on the implementation of the pro-posed model.

ACKNOWLEDGMENTS

This work is supported in part by the U.S. National Science Foundation (NSF) under grant HRD 1828811. However, any opinion, finding, and conclusions or recommendations expressed in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the NSF.

REFERENCES

[1]. Mell, Peter, Grance, Tim. The NIST Definition of Cloud Computing Version 15, 10-7-09 National Institute of Standards and Technology, Information Technology Laboratory, 2011

[2]. Magnabosco, John . Protecting SQL Server Data, 2009

[3]. Rawat, Danda B. and Ghafoor, Kayhan Zrar, "Smart Cities Cybersecurity and Privacy", Elsevier Press, ISBN: 9780128150320, Dec. 2018

[4]. S. Jeschke, C. Brecher, H. Song and Rawat, Danda B. (Eds.), "Industrial Internet of Things: Cyber-manufacturing Systems," ISBN 978-3-319-42559-7, Springer, Switzerland, 2016.

[5]. Identity Theft Resource Center [ITRC) Annual Data breach Report, 2017

[6]. Identity force , Data-breaches, 2017

[7]. Kumar,Vimal, Chaisiri, Sivado and Ko , Ryan “Data Secu rity in CloudComputing” The Institution of Engineering and Technology, 2017

[8]. Ristenpart, Thomas , Tromer, Eran , Shacham, Hovav ,and Savage,Stefan. Hey, you, get off of my cloud: Exploring information leakage in third-party compute clouds. In Proceedings of the 1 6th ACM conference on Computer and communications security, pages 199–212. ACM, 2009.

[9]. Chen, Shuo , Wang, Rui , Wang, XiaoFeng and Zhang, Kehuan . “Side-channel leaks in web applications: A reality today, a challenge tomorrow” IEEE Symposium on Security & Privacy, pages 191–206, 2010

[10]. Suzaki, Kuniyasu , Iijima, Kengo, Yagi, Toshiki and Artho, Cyrille. “Memory deduplication as a threat to the guest os. In Proceedings of the Fourth European Workshop on System Security, page 1. ACM, 2011.

[11]. Janacek, Bob.”Best Practices: Securing Data at Rest, in Use, and in Motion” ,2015

[12]. M. A. AlZain, B. Soh, and E. Pardede, “Mcdb: using multi-clouds to ensure security in cloud computing,” in Dependable, Autonomic and Secure Computing (DASC), 2011 IEEE Ninth International Conference on. IEEE, 2011, pp. 784–791.

[13]. Asaf Shabtai,Yuval Elo vici, and Lior Rokach. A Survey of Data Leakage Detection and Prevention Solutions. Spring er Science & Business Media, 2012.

[14]. Nyarko,Richard “ Security of Big Data: Focus on Data Leakage Prevention (DLP)” , 2018

[15]. Begna, Geremew Rawat, Danda B., Garuba, Moses and Njilla, Laurent "SecureCASH: Securing Context-Aware Distributed Storage and Query Processing in Hybrid Cloud Framework," Proc. of the 2018 IEEE Conference on Communications & Network Security (IEEE CNS 2018)-Posters, Beijing, China, May 2018.

[16]. Begna, Geremew ‘a web-based approach to role-based access control, may, 2010

[17]. N. M. Smith, W. C. DeLeeuw, and T. G. Willis, “Diffie-Hellman key agreement using an M-of-N threshold scheme,” Jan. 2 2018. US Patent 9,860,057

[18]. "Math Forum: Ask Dr. Math FAQ: The Birthday Problem". Mathforum.org. Retrieved 20 October 2018

[19]. Evans , Michael J. and Rosenthal,Jeffrey S. Probability and Statistics The Science of Uncertainty ,Second Edition University of Toronto