Journal of Technology and Systems

Vol. 3, Issue No. 1, pp 14 – 21, 2021

www.carijournals.org

DATA PROTECTION TECHNIQUES: A CRYPTOGRAPHY APPROACH

[1*]Laszlo Reynolds

[1]Postgraduate Student Finstock Evarsity College

[*]Corresponding Author's Email: laszloallan@gmail.com

[2*]Johnson Winks

Lecturer: Finstock Evarsity College

[*]Corresponding Author's Email: journals@carijournals.org

Abstract

Purpose: Digital world is a complex and a constantly evolving world alongside the modernization. This has been through the improvement of the technology from physical nature information to digitized data that is the information technology. Data is, therefore, becoming a basic element in today's world than they were anticipated some three decades ago, now being shared and stored electronically. This study, therefore, sought to establish the ways in protecting information privacy using cryptography.

Methodology: The study adopted a desktop literature review method (desk study). This involved an in-depth review of studies related to data protection and cryptography. Three sorting stages were implemented on the subject under study that is data protection and cryptography in order to determine the viability of the subject for research. The first stage that comprised the initial identification of all articles that were based on data protection and cryptography from various data bases. A second search involved fully available publications on the subject of data protection and cryptography. the third step involved the selection of fully accessible publications. Reduction of the literature to only fully accessible publications yielded specificity and allowed the researcher to focus on the articles that related to data protection and cryptography which was split into top key words. After an in-depth search into the top key words (data protection and cryptography), the researcher arrived at 12 articles that were suitable for analysis. Analysis was done using Excel where the study presented the findings in form of themes.

Findings: The research findings point towards several methods that can be used in cryptography which were concluded to include symmetric encryption, asymmetric encryption and hashing.

Unique contribution to theory, policy, and practice: Previous studies were reviewed and some of them presented several knowledge gaps. The review of the literature presented knowledge gaps in the contextual, conceptual and methodological fronts. For instance, one of the gap that was presented was the contextual gap (geographical gap). This is because some of the studies were done in different geographical contexts such as the United States, China and Japan. The application of the findings will be challenged by the technological advancements in Kenya since Kenya is still a developing country in the process of up taking new technologies. Other studies in their research applied different methodologies such as the use of the FTK Imager and

the use of Hashing. The studies did not entirely focus on the use of cryptography as a method of data security and thus they presented a methodological gap.

Keywords: *Data Protection, Cryptography.*

1.0 INTRODUCTION

1.1 Background of the Study

The development of computer science has exponentially, progressed and advanced over the past decades to elevate the information technology to a digital world where services and infrastructures are electronically being operated and run-on numerous versions. This development has been appreciated by the new generations due to its ability to improve operations alongside the modernization (that is from physical nature to digitized nature of information technology). Digital (electronic) data is, therefore, becoming a basic element in today's world than they were anticipated some three decades ago, now being shared and stored electronically (Austin, 2016).

With the rapid advancement in technology, and globalization of businesses/multinationals, significant amount of data is constantly being transacted through the internet (Oetzel & Miklian, 2017). Digital information is to some extent very fragile and volatile in nature especially when it comes from purportedly illegal, illicit and malicious activities. Privacy, therefore, is very essential in order to protect the integrity of the users. Confidentiality on the other hand acknowledges the treatment of identifiable information that has been disclosed to others in relation of trust and with the expectation that it will not be divulged to others except in previously agreed-upon ways (Khan, 2017).

The automated arrangement, coordination, and management of complex computer systems, and services through the internet and digital devices denotes that every person is connected by means of the wealth of online services that encompass all the sectors of the contemporary world. This, therefore means that everyone will have to use some form of digital information or electronic services to fulfil their daily routine (Goel & Dey, 2016).

However, the explosion of the internet, and with the advancements and innovations to social media platforms such as Email, Facebook, twitter, WhatsApp, etc. which has been critical to the well-functioning of governmental, business, educational, social, entertainment institutions have also dragged along the channels and loopholes for abuse and crimes against the creative power of these digital resources. Qualified as cybercrimes, these abuse and intrusions such as malware, identity thefts, large scale digital fraud and embezzlement, human and illicit materials trafficking, and violation of Integrated Population Registration Services (IPRS) have dramatically increased (Ablon, 2018).

This has thus necessitated the need for such information to be kept within the confines of the designated purpose and users resulting to the information security (Saleem, Popov & Dahman, 2011). Security measures against these digital crimes have thus been set up from individual operating procedures which limit the search for evidence to the development of forensic investigation which requires protection of the collected evidence against potential attacks which may take advantage of the volatile nature of digital evidence (Nieto, Rios & Lopez, 2018).

One of the mechanisms is the encryption of data. In the presence of a third party in the chain of operation involving digital information, cryptography comes in handy to ensure privacy. Cryptography is the science of using mathematics to encrypt and decrypt data to keep messages secured by transforming intelligible data form (plaintext) into unintelligible form (cipher text) (Saleh, Aly & Omara, 2016).

Encryption of personal data is widely regarded as a privacy preserving technology which could potentially play a key role for the compliance of innovative IT technology (Spindler & Schmechel, 2016). This technique is becoming increasingly essential for many enterprises that are data-driven and for preserving data subjects' privacy with regard to today's monitoring and profiling possibilities (both in the government institutions as well as high-tech companies).

1.2 Research Problem

Amongst the many protection techniques that are existing in the world today, not all of them have been able to offer an all-round protection solution that can be long term. Processes are in the rapid development every day considering the advancing technology and the vigorous competition between businesses to survive the shifts. Amongst the numerous techniques to employ, an individual/company will choose the most effective method that offers reliable and long-term protection to their data.

According to Saleh, Aly and Omara (2016), in using cryptography, the cipher text looks meaningless, so the attacker will interrupt the transmission or make more careful checks on the data from the sender to the receiver. Steganography as per Douglas, Bailey, Leeney and Curran (2018) can be described as the art and science of covert communications which involves the process of hiding information inside other information. However, the challenge with steganography is that once the presence of hidden information is revealed or even suspected, the message becomes known to a third party (Spindler & Schmechel, 2016). On the other hand, the use of digital signatures as a technique in securing digital data is not enough. It does not provide confidentiality of the data although it is in an encryption format. The user has to ensure that the private key used is inaccessible to a third party (Poonguzhali & Moorthy, 2017).

Despite cryptography being the best solution to information privacy protection, there still exists a contention as to the transmission and storage of the encryption keys used. In cryptography, there are two key issues. One of them is how to transmit the encryption key to the intended user in a secure connection. The other issues arise where, given a large number of keys, how will the user be able to store and memorize the keys without having them being disclosed to unintended parties/users. Therefore, it is against these backgrounds that the current study seeks to establish a technique for protecting information privacy using cryptography.

1.3 Main Objective

To establish approaches for protecting information privacy using cryptography.

2.0 LITERATURER REVIEW

2.2 Theoretical Review

This study was informed by the diffusion theory and the technology acceptance theory.

2.2.1 Diffusion of Innovation Theory

The theory was developed by a scholar known as Rogers who contends that diffusion is the process by which an innovation is communicated over time among the participants in a social system. For Rogers (2003), adoption is a decision to utilize innovation fully as the available best course of action and rejecting any decision not to embrace the technology. It is the process by which an innovation is channeled over time among the members of a social system.

Wesley (2018) relates the diffusion innovation to a renowned concept called cryptocurrency innovation. It is a digital / virtual currency that uses cryptography for security, which is difficult to counterfeit because of this security feature. His approach sought to qualitatively collect primary and secondary data focus on collecting data from a range of different countries allowing for a wider understanding of the phenomenon. He unveiled a wide range of trends and developments in the diffusion process.

Woodside, Augustine & Giberson (2017) likewise are in acceptance of the theory and their focus was on the acceptance and future use of blockchain technology. They looked at how the technology has been implemented for instance in the use of bitcoin. A cryptocurrency also applies here, by allowing a medium of exchange similar to the US dollar, though it is digital and utilized encryption to control new currency creation and verification of funds (PwC, 2016; Capgemini, 2017). Blockchain technology uses peer-to-peer networking without the need for a centralized server, and instead the blockchain exists across an entire network of computers (Lord, 2016). This theory, thus, informs the study to seek to enhance on the cryptography transmission mechanism in consideration of the advancements in technology.

2.2.2 Technology Acceptance Model (TAM)

TAM was introduced by Fred Davis in 1986 specifically tailored for modeling users' acceptance of information systems or technologies. That is, the theory is used to explain the general determinants of computer acceptance that lead to explaining users' behaviour across a broad range of end-user computing technologies and user populations. The basic TAM model included and tested two specific beliefs: Perceived Usefulness (PU) and Perceived Ease of Use (PEU). This model is widely used to study user acceptance of technology. According to TAM, perceived usefulness (PU) and perceived ease of use (PEU) influence one's attitude towards system usage, which influences one's behavioural intention to use a system, which, in turn, determines actual system usage.

Perceived Usefulness (PU) as 'the degree to which a person believes that using a particular system would enhance his or her job performance (Davis, 1989) and PEU as 'the degree to which a person believes that using a particular system would be free of effort. Perceived ease of use is predicted to influence perceived usefulness, because the easier a system is to use, the more useful it can be. These constructs reflect users' subjective assessments of a system, which may or may not be representative of objective reality. System acceptance will suffer if users do not perceive a system as useful and easy to use (Lai, 2017).

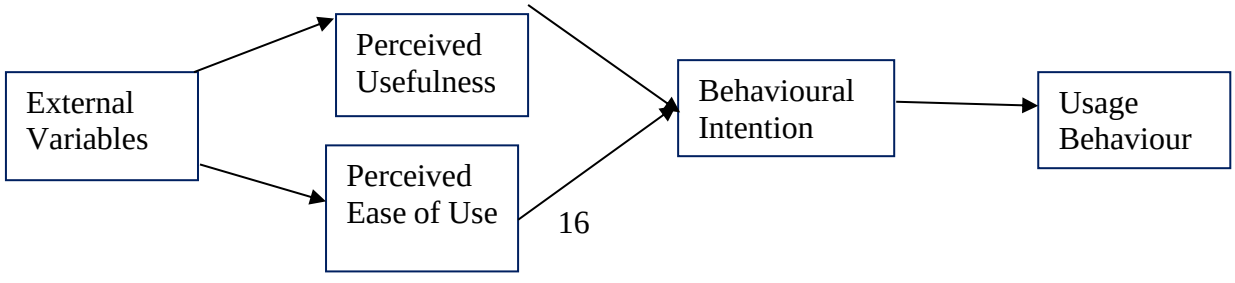

Figure 1: Technology Acceptance Model (TAM)

The model has been advanced and used in many business and entrepreneurial contexts such as the health care, education, banking and public service sectors. The emergence of mobile payment has enabled customers to embrace E-commerce with the use of mobile transactions. Which involves online shopping, bookings, e-tickets, etc. The M-commerce websites are collaborating with payment providers such as credit/debit card providers to link to a third party provider such as Paypal to securely accept payment from customers through cryptography authentication tokens Known as JWT (Json Web Tokens) or OAuth2.0 Tokens. The emergence of M-wallet allows the user to easily download and install the application on the smartphone and thus it is able to store the customer's financial information in the SIMcard or securely in cloud. The use of cryptography allows such transactions to be personalized and secure personal information (Madan, 2016; Lee, 2017). The model is very instrumental to the study since it lays a foundation on which the cryptography model is built upon. It is from the theory that the study therefore looks to theoretically fill in the gap that exists between the transmission of cryptography key in a secure manner.

2.3 Conceptual Framework

Smith (2004) defines a conceptual framework as a hypothesized model identifying the framework under study and showing the relationship between the variables in the study i.e. dependent and independent variables. It is a research tool intended to assist a researcher to develop awareness and understanding of the variables. For the purpose of this research, a conceptual framework has been developed showing the relationship between the independent variables and the dependent variable. The framework postulates a relationship of independence and dependence that is to come up with a model and then try by reality testing to see if the relationships actually work out that way (Yamauchi, Ponte, Ratliffe & Traynor, 2017).

The dependent variable in this study, will be measured by the use of cryptography. The independent variable is denoted by cryptography as a technique in data protection while the dependent is derived as the protection of information.

The current study, therefore, seeks to focus and apply the Role-Based Access Control (RBAC) and the notification techniques in establishing their effectiveness in securing digital data evidence. This will, therefore be done using the following proposed model:

2.3.1 Conceptual model

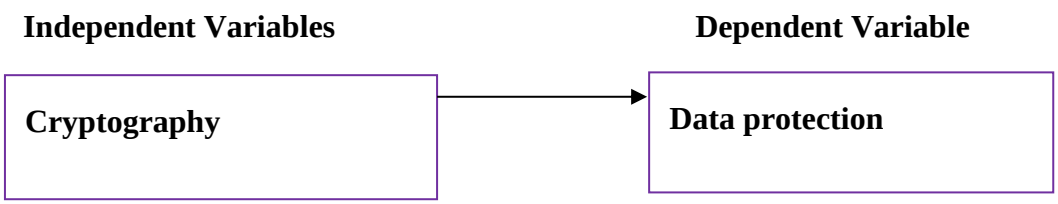

Figure 2: Conceptual Framework

3.0 METHODOLOGY

The study adopted a desktop literature review method (desk study). This involved an in-depth review of studies related to data protection and cryptography. Three sorting stages were implemented on the subject under study that is data protection and cryptography in order to determine the viability of the subject for research. The first stage that comprised the initial identification of all articles that were based on data protection and cryptography from various data bases. The first search was done generally by searching the articles in the Article title, abstract, keywords. A second search involved fully available publications on the subject of data protection and cryptography. A total of studies that were studied totaled to 50 articles. The Web of Science data base was not accessible to the researcher due to limitations in the rights to the database and as such the researcher opted to not pursue that direction. The filtration process was done basing on the currency of the articles (between the year 2014 and 2019). This section indicates the process that the study followed in analysis of the literature. The researcher reviewed the articles to eliminate duplicates; this ensured only unique studies for unique review. The study took an approach of English literature since it is common globally; thus, the focus was on data protection and cryptography in English language. The selected literature was then sorted and only journals were included. This was done according to the **ABDC** and **ABS** lists. After the filtration was done, the third step involved the selection of fully accessible publications. Reduction of the literature to only fully accessible publications yielded specificity and allowed the researcher to focus on the articles that related to data protection and cryptography which was split into top key words. After an in-depth search into the top key words (data protection and cryptography), the researcher arrived at 12 articles that were suitable for analysis. Analysis was done using Excel where the study presented the findings in form of themes.

4.0 FINDINGS AND PRESENTATION

4.1 Cryptography

The main mechanism applied in cryptography is data encryption. Encryption takes a piece of data, commonly called the plaintext, together with a cryptographic key and produces a scrambled version of the data called the ciphertext. Using the key, it is possible to decrypt the data to recover the plaintext, but without the key the ciphertext hides all information about the original data, other than its length. This security property, commonly known as semantic security guarantees that, without the key, an adversary cannot learn any (potentially sensitive) property of the underlying data even if he has a lot of insight as to what the data may be. This is critical in applications where data may have some predefined structure, such as in financial transactions or if partial information about the underlying distribution of data is known, such as when the data is measuring some real-world phenomenon (Izhar, Kaushal, Fatima & Qadeer, 2017).

The encryption consists of protocols as shown below:

- KeyGen - a key generation algorithm that generates the necessary cryptographic keys,
- Enc(k, p) = c - an encryption algorithm that uses a key k to scramble the plaintext p into cipher text c,
- Dec(k, c) = p - a decryption algorithm that uses the key k to recover the plaintext p

from the ciphertext c.

There are two methods of encryption: symmetric and asymmetric encryption.

4.2 Symmetric encryption

According to Stallings (2017), symmetric encryption (secret key encryption) is effective when the sender and the recipient possess the same keys to encrypt and decrypt a message. Both keys must be kept secret. In case one of the keys is compromised, further secure communication is impossible. Keys need to be exchanged between users, often over a slow secure channel, for example a private courier. However, the number of keys can be very large, if every pair of users requires a different key, even for a moderate number of users, i.e., n (n — l)/2 for n users. This creates a key-distribution problem which is partially solved in the asymmetric systems. Examples of symmetric systems are the data encryption standard (DES) and rotor ciphers (Khan, 2017).

In the case of symmetric system of many users (n), each time a new user is added to the system, it needs only a public key and a private key. Thus, for n users, we have 2 n keys, which is O (n). However, it will depend on the type of the algorithm and each user may require separate pairs for confidentiality and signing, i.e., 4 n keys, which is still O (n) (Gong, 2017).

Symmetric key cryptography has the advantages of fast speed and high security. However, the key cannot be secretly allocated. In addition, it lacks of the ability to automatically detect the key leak. In a situation where the two users operate two different networks, the users have to have separate keys for the different networks that is the number of keys will automatically increase.

4.3 Asymmetric encryption

On the other hand, asymmetric encryption also referred to as public key encryption uses what is called a key pair (a public key for encrypting a message, and a private key to decrypt it) (Stallings, 2017). In the existence of practical problems of distribution of a large number of keys, a solution to this problem was suggested by Diffie and Hellman (1976). The proposed type of cipher involved 2 separate keys where one is used for enciphering which can be made public, while the deciphering is carried out by the other key and thus, this second key is private.

For instance, if user M intends to send a message to N, M can use N's public key to encipher the data. Only N can decipher the ciphertext because he has the secret deciphering key. This is what ciphers Khan (2017) refers to as public-key cryptosystem or an asymmetric cryptosystem. If asymmetric algorithms satisfy certain restrictions, they can also be used for generating so-called digital signatures. Examples of Asymmetric systems are ElGamal, Diffie–Hellman key exchange, RSA etc.

In asymmetric encryption, a model representation can be used to explain further. Each subject S has a publicly disclosed key K_s where s is the public key that anyone can use to encrypt, and a privately held key K_s^{-1} where s is the private key. The association can thus be represented as follows:

$$M = \{\{M\}_{K_S}\}_{K_S^{-1}}$$

It thus noted that only the sender denoted by K_s^{-1} is the one able to decrypt this information. In a situation where there is more than one user (n), each time a new user is added to the system, the system is required to share a new key (a public and a private key) with to the new user including the previous ones. A model representation would look like this: $1 + 2 + \ldots + (n - 1) = n (n - 1) /2$ keys. This is O (n^2) keys.

Asymmetric key encryption offers the benefit of open encryption keys, users do not need to pass between the private key. One can also achieve the digital signature with ease to ensure the confidentiality of the transaction process, identity verification, data integrity. However in asymmetric key encryption the technology is very complexity which makes the decryption process a little slower.

4.4 Hashing

Hash functions are one-way encryption algorithms that to some extent do not require the use of keys. It will accept a variable-length block of data M as input and produces a fixed-size hash value h = H (M) that is the main agenda / view point of hashing is the integrity of data. Any detection in the change of a function in the hash, results in high likelihood of alteration in the code of the hash. Virtually all cryptographic hash functions involve the iterative use of a compression function. A cryptographic hash function can be used to construct a pseudorandom function (PRF) or a pseudorandom number generator (PRNG). Secure Hash Algorithm (SHA) is a family of cryptographic hash functions (Poonguzhali & Moorthy, 2017).

Shah, Saleem and Zulqarnain (2017) resonate with the methodology of the technique. The study extrapolates on hashing by the use of a protection tool known as FTK Imager. It supports four different formats to store the extracted image. These formats are AD1, E01, RAW and SMART. The integrity of digital evidence is guaranteed by ascertaining MD5 and SHA1 hashes of the extricated substance and then stored in a report alongside different points of interest identified with the drive. It additionally offers an encryption highlight to guarantee the secrecy of the computerized evidence. The evidence can be encrypted by utilizing a secret passcode or a computerized certificate.

However, according to the authors, FTK Imager has some few shortcomings, for instance, it lacks a function to verify / authenticate the user who is extracting the forensic image. There is no way of knowing that evidence was, indeed, extracted by an authorized person. So, there is a big question on the soundness of the evidence. Secondly, integrity of the evidence is provided through hashes and hashes prove to be not guarantee enough integrity of the evidence.

Wang (2017) likewise advises that storage of user passwords in a hash digest could be an effective way of ensuring security of information. The passcode presented by the user is then hashed and a comparison with the one stored is done in order to ensure the user is genuine. During this process a passcode reset mechanism is also required and the original ones cannot be duplicated from the password stored.

5.0 CONCLUSIONS

Previous studies were reviewed and some of them presented several knowledge gaps. The review of the literature presented knowledge gaps in the contextual, conceptual and methodological fronts. For instance, one of the gap that was presented was the contextual gap (geographical gap). This is because some of the studies were done in different geographical contexts such as the United States, China and Japan. The application of the findings will be

challenged by the technological advancements in Kenya since Kenya is still a developing country in the process of up taking new technologies. Other studies in their research applied different methodologies such as the use of the FTK Imager and the use of Hashing. The studies did not entirely focus on the use of cryptography as a method of data security and thus they presented a methodological gap.

REFERENCES

Ablon, L. (2018). *Data Thieves: The Motivations of Cyber Threat Actors and Their Use and Monetization of Stolen Data*. RAND.

Austin, T. (2016). Towards a digital infrastructure for engineering materials data. *Materials Discovery*, *3*, 1-12.

Capgemini. (2017). A History of Bitcoin. Retrieved from https://www.capgemini.com/beyond-the-buzz/cryptocurrency-blockchain

Douglas, M., Bailey, K., Leeney, M., & Curran, K. (2018). An overview of steganography techniques applied to the protection of biometric data. *Multimedia Tools and Applications*, *77*(13), 17333-17373.

Goel, A., & Dey, S. (2016). *A fingerprint based crypto-biometric system for secure communication*. Discipline of Computer Science and Engineering, IIT Indore.

Izhar, S., Kaushal, A., Fatima, R., & Qadeer, M. A. (2017, November). Enhancement in data security using cryptography and compression. In *2017 7th International Conference on Communication Systems and Network Technologies (CSNT)*(pp. 212-215). IEEE.

Khan, M. T. (2017). Network Security Mechanisms and Cryptography. *International Journal of Computer Science and Mobile Computing*, (6)7, 138-146.

Lee, I. (Ed.). (2017). *The Internet of Things in the Modern Business Environment*. IGI Global.

Lord, S. (2016). Bankchain & Itbit: Settling On The Blockchain. Modern Trader, pp 16-21.

Madan, S. (Ed.). (2016). *Securing Transactions and Payment Systems for M-commerce*. IGI Global.

Oetzel, J., & Miklian, J. (2017). Multinational enterprises, risk management, and the business and economics of peace. *Multinational Business Review*, *25*(4), 270-286.

Poonguzhali, P. K., & Moorthy, N. A. (2017). Design of a Dynamic Clustering With Secured Hashing Technique in Wireless Sensor Network. *Asian Journal of Applied Science and Technology (AJAST)*, *1*(4), 28-33.

PwC. (2016). Making sense of bitcoin, cryptocurrency, and blockchain. Retrieved from https://www.pwc.com/us/en/financial-services/fintech/bitcoinblockchain-cryptocurrency.html

Saleh, M. E., Aly, A. A., & Omara, F. A. (2016). Data Security Using Cryptography and Steganography Techniques. *IJACSA) International Journal of Advanced Computer Science and Applications*, *7*(6), 390-397.

Shah, M. S. M. B., Saleem, S., & Zulqarnain, R. (2017). Protecting digital evidence integrity and preserving chain of custody. *Journal of Digital Forensics, Security and Law*, *12*(2), 12.

Spindler, G., & Schmechel, P. (2016). Personal data and encryption in the European general data protection regulation. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *7*, 163.

Stallings, W. (2017). *Cryptography and network security: principles and practice* (pp. 92-95). Upper Saddle River: Pearson.

Wesley, M. (2018). An investigation into the diffusion of the cryptocurrency innovation.

Woodside, J. M., Augustine Jr, F. K., & Giberson, W. (2017). Blockchain technology adoption status and strategies. *Journal of International Technology and Information Management*, *26*(2), 65-93.

Yamauchi, L. A., Ponte, E., Ratliffe, K. T., & Traynor, K. (2017). Theoretical and Conceptual Frameworks Used in Research on Family-School Partnerships. *School Community Journal*, *27*(2), 9-34.