Asian Journal of Applied Science and Technology
Volume 4, Issue 3, Pages 119-128, July-September 2020

A Brief Survey on Various Technologies Involved in Cloud Computing Security

Vindhuja E.[1] & Dr.N.Umadevi[2]

[1]*M.Phil Research Scholar, Department of Computer Science, Sri Jayendra Saraswathy Maha Vidyalaya College of Arts and Science (Affiliated to Bharathiar University), Coimbatore-641005, Tamil Nadu, India.*
[2]*Associate Professor & Head, Department of Computer Science, Sri Jayendra Saraswathy Maha Vidyalaya College of Arts and Science, (Affiliated to Bharathiar University), Coimbatore-641005, Tamil Nadu, India.*

DOI: 10.38177/ajast.2020.4316

Article Received: 07 June 2020 *Article Accepted: 21 August 2020* *Article Published: 22 September 2020*

ABSTRACT

In the past decade, big technical advances have appeared which can bring more comfort not only in the corporate sector but at the personal level of everyday life activities. The growth and deployment of cloud computing technologies by either private or public sectors were important. Recently it became apparent to many organizations and businesses that their workloads were moved to the cloud. However, protection for cloud providers focused on Internet connectivity is a major problem, leaving it vulnerable to numerous attacks. Although cloud storage protection mechanisms are being introduced in recent years. However, cloud protection remains a major concern. This survey paper tackles this problem by recent technology that enables confidentiality conscious outsourcing of the data to public cloud storage and analysis of sensitive data. In specific, as an advancement, we explore outsourced data strategies focused on data splitting, anonymization and cryptographic methods. We then compare these approaches for operations assisted by accuracy, overheads, masked outsourced data and data processing implications. Finally, we recognize excellent solutions to these cloud security issues.

Keywords: Cloud computing, Security, Data splitting, Anonymization, Cryptographic techniques.

1. Introduction

In certain ways, cloud computing has become widely used, such as file sharing, real-time applications and communications. Big advances in cloud computing, including major development, appeared in recent decades. The practical deployment of these advancements, which can theoretically provide functionality on many fronts, has made cloud storage commonly accessible both in private and public sectors. The reliability of the services offered is very efficient at the same time numerous problems were occurred both for cloud customers and cloud service providers. Security in cloud storage is a crucial data safety subdomain and constitutes a significant barrier to the widespread implementation of cloud technology [1]. Because cloud-based systems simply use an Internet connection, they are open to a range of attacks and other security risks that can lead to serious consequences, including data leaks, malware intrusion, Denial-of-Service (DoS) attacks, lack of data and unsecured APIs [2]. According to [3], cases of vulnerability in the cloud world have increased in recent years mostly due to the significant rise in cloud resources.

It has become apparent that several organizations and businesses have increasingly started to move their workloads to the cloud. The number of businesses switching to the cloud will rise to 83 percent by the end of 2020, according to a survey conducted by Logic Monitor [4] which is a leading SaaS-based performance management tool.

Despite the advancements in cloud use, the implementation of this omnipresent infrastructure will face many major challenges. The study by a leading cloud computing firm, Right Scale [5], reports that the cloud protection issue continues to constitute a significant problem, along with increased investment, lack of infrastructure and knowledge and efficiency problems (including others). For this article, we performed a detailed survey to look at different forms of possible defense approaches associated with cloud computing security issues. The rest of this article has organized in the following. Section 2 covers similar studies in recent years on cloud security. In Section

3 we address different methodologies used with their advantages and drawbacks in cloud computing security. In Section 4 we examine the comparative results for the surveyed methods and in Section 5 we conclude the article.

2. Related Works

In 2015 the author [6] had outlined the key problems for cloud computing security and privacy, listed current technologies and contrasted their benefits and drawbacks, but their analysis papers did not compare them to many of them. In 2016 the author [7] had proposed hidden communication techniques including hidden channel attacks and fuzzy-based systems. They had mentioned their benefits and weaknesses. The hidden communication technology was applied to a real-time scenario. But somehow its lack of accuracy level. In 2017 the author [8] proposed a related user access method in access management, which is a critical tool to ensure information safety, fairness, and privacy within the network protection can be made possible by a series of regulations and procedures.

In 2018 the author [9] analyzes the definition and the functionality of edge computing. They have been implemented in a cloud-based scenario where some criteria have been presented for edge computing technology data analysis gives an overview of the possible security dangers of edge computing and an exhaustive study of current edge computing data analytics briefs with its advantages and drawbacks. In 2019 the author [10] analyzed many blockchain-based protection approaches that included verification, encryption, data and resource access lists with integrity assurance. But other facets of data security technologies, such as trust, credibility and access management, were missing.

3. Methodologies

This segment discusses the state-of-the-art technologies which retain the data security strategies that were used as data protection measures in cloud computing. Our research concentrates on the security measures which were maintained in the cloud computing also with their advantages and disadvantages. The mechanism that we had taken in our survey are:

(i) Data splitting mechanisms

(ii) Data anonymization methods and

(iii) Cryptographic techniques.

The first two are indirect security mechanisms in cloud computing where many researchers are not concentrated on it but we had taken this to create a novel survey. As can be seen in our comparative study for the above mechanism we had taken some parameters to show the effectiveness of each security mechanism.

(i) Protecting the capacity for each mechanism

(ii) Level of security accuracy

(iii) Computation Overhead

(iv) Different Functionalities

(v) Major essential criteria for handling security by each mechanism.

3.1 Data Splitting Method

The data splitting is a method of security focused on the separation of confidential data and the storing that collection in different locations as fragments. Fragments should be in the form in which it does not cause the subject to be re-identified, nor does it expose sensitive information relating to a specific item. For instance, where a fragment consists of a 'diagnosis' attribute values, it is pointless for an attacker to know explicitly a list of diagnoses since they cannot be paired with the appropriate individuals. Data may be outsourced in the cloud by a local proxy that separates data into the same CSP, or multiple clouds, each running a separate CSP that provides the same type of service (i.e. a multi-site) in the Cloud scenario. To ensure efficient data security based on separation, storage sites can stay unconnected and separate, so that CSPs cannot join forces with partial data collection in the hope of breaching data privacy.

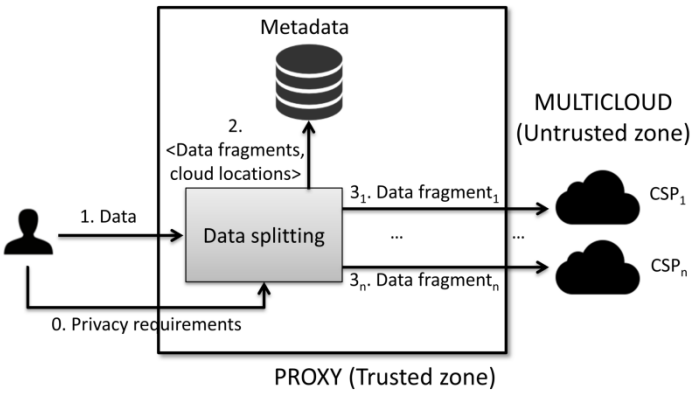

Fig.3.1 Data Splitting Workflow In Cloud Computing

Figure 3.1 illustrates the method of separating and storing of the data. First, the proxy collected the data to be externalized from the customer and analyzed the disclosure risks. To do this, the proxy depends on the user-defined privacy criteria (stage 0 in Figure 3.1), which describe a set of attributes that may re-identify, i.e. the identity or quasi-identification attributes. The proxy decides how the data is distributed and how many storage places to avoid leakage. A data fragment is given by some piece of data that can be placed securely together. The framework also stores the divisional criteria and storage positions in a local database (Step 2), so that subsequent searches of divided data can be interpreted correctly and the partial results can be applied. Finally, each data fragment is transmitted to a different CSP (Steps 3_1 to 3_n). In this way the method of breaking up is loss-free (thus, the findings that will be collected on the initial data set can be replicated without needing to retain this data on the local premises) and confidential (as the fragments stored on a different CSP are divulgated in compliance with the conditions for confidentiality).

Advantages

- Because of the simple handling of the fragments estimates can be easily assisted on individual fragments.
- However, data calculations stored in various clouds involves some algorithms that break and orchestrate the computations between the CSPs privacy-protected by the local proxy.

- CSPs often maintain maximum functionality for storing data fragments on a fragment that they store that is part of the use of the entire dataset.

Disadvantages

- Remember that, if many users have one local proxy and thus share one cloud locations, pieces of users' data can be stored at each location, meaning fragmented users' data at each cloud location are fused.
- Another drawback is the fact that the proxy is the only one able to understand the fragment is the person who offers defense against CSP collusion. Although the CSP will capture and monitor the partial data they hold, the individual to which each fragment belongs cannot undeniably be established. This introduces a single fault point, however, as cloud data cannot be retrieved and fragments cannot be recomposed without the metadata stored by the proxy.
- The proxy's metadata replication and recovery procedures are also essential to guarantee the mechanism's robustness but create many duplications.

3.2 Data Anonymization Method

Data anonymization irreversibly hides the data in a privacy-preserving manner, in contrast to data separation (and even data encryption). To expose sensitive data to an untrusted third person, anonymization strategies have been developed within the field of predictive disclosure management and secrecy data protection, so that they are analytically useful for secondary use and do not expose details that may be attributed to particular persons.

The untrusted entity, in our case, is the CSP and the secondary apps are the outsourced calculations on masked data that consumers (e.g. data analysts) want to run. Alternatively, anonymized data preserve its value for anyone while encrypted data is worthless to everyone who does not decode. CSPs, in particular, who gratuitously provide their services for the monetization of the data analyzes they maintain, can object to encrypted data (unless the encryption takes place by themselves), but not to anonymized data (in the majority of cases their software for automatic processing cannot even detect this anonymization) which was shown in Figure 3.2.

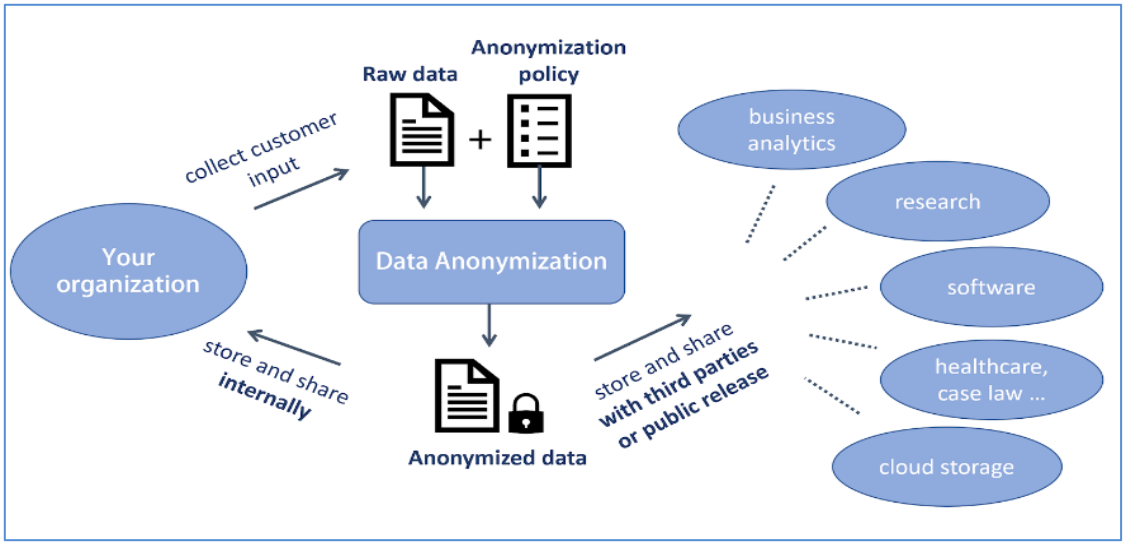

Fig.3.2 Data Anonymization Workflow in Cloud Computing

For each document or collection of records, anonymization may be done separately. In the first example, anonymized data changes can be enabled quickly, but in the second case, changes require a re-anonymization of the affected database category or the whole data collection. Although anonymization can be required on all the data set attributes to be outsourced, anonymizing the subset of attributes that will potentially require a leak is adequate throughout most cases.

Attributes can be categorized as Depending on their transparency capacity as follows:

- **Identifiers:** The subject's name (i.e. social security numbers) is revealed separately. These attributes must either be omitted or replaced by random values from the anonymized data set.
- **Quasi-identifier or main attributes:** In contrast to identity, such attributes do not classify particular objects, but can combine them. For example, for a certain age, gender, employment, or zip code, there may not be one single person, but for similar age, gender, job and zip code mix (see a 95 years old female doctor living in a poor neighborhood), who would be easy to find. Consequently, anonymity should be granted to almost identity attributes.
- **Confidential characteristics:** They express a person's sensitive characteristics (income, sexuality, health status, etc.). Where sensitive characteristics cannot be paired with a name, they are unmodified and are useful for empirical use. In certain cases, however, they are anonymized to gain better protection. In addition to quasi-identifiers
- **Non-confidential:** Features not belonging to any of the above groups remain unchanged.

Advantages

- However, anonymized data has the big advantage of being more conveniently accessed relative to encrypted data and data storage.
- Masking is carried out only at the storage stage for anonymous data, and it can be done for linear or quasi-linear algorithms any query for anonymized data (search, recuperation, calculation) is straightforward and would not incur overhead for either the proxy or the CSP.
- The irreversible confidentiality of the data does not hide (in essence they cannot) the data collected and the measurement results, which ensures that local proxy is needed only at the storage stage and there is no need to store or handle the main content.

Disadvantages

- Confidential statistics were represented from the usually estimated data rather than the precise outcomes of estimates.
- A second drawback is the lack of support for keyword searches on anonymized attributes. In reality, when anonymization of data modifies individual values but helps to preserve the integrity of the entire data collection, this is ideal for statistical comparisons (for example, statistics) of outsourced data sets, just not

for questions or evaluations of some persons. Although inferences from anonymized data were possible to draw from such people, there would be no anonymity.

3.3 Cryptographic Techniques

In several of the technology strategies implemented to resolve privacy issues in the cloud, cryptography can be used as one of the best and most important building blocks. Modern cryptography has developed to meet different security needs in the real world, many of them in the cloud market. Cryptography requires a wide variety of information management priorities, interfaces and features into account. Cloud scenarios in real-world which involve different security reasons, such as protection of data or data transparency. Also, features may be more robust than mere secure cloud storage, such as cloud search and retrieval of data or processing of secured data.

As the data breach epidemic in the past decades has been shown, many cloud storage providers are now using cryptographic solutions, especially via symmetric encryption schemes such as AES256, to encrypt data they hold. However, the core materials are also managed by the CSP in these systems, and the protection of the outsourced data relies on the trusting relationship between consumers and the CSP.

The data is encrypted using conventional cryptographic methods before outsourcing, as normal cryptography to further safeguard privacy toward ineffective CSPs. This strategy however forbids CSPs from carrying out any sophisticated programming and requires users to import the whole data collection if the outsourced applications are to be used. Consequently, standard encryption systems do not allow the transition of the bulk of information to the cloud through the outsourced encrypted data, and this significantly limits server technical efficiency and savings.

Recent encryption research focuses on improving encryption strategies that allow customers to safely transfer sensitive features on outsourced encrypted data to the cloud.

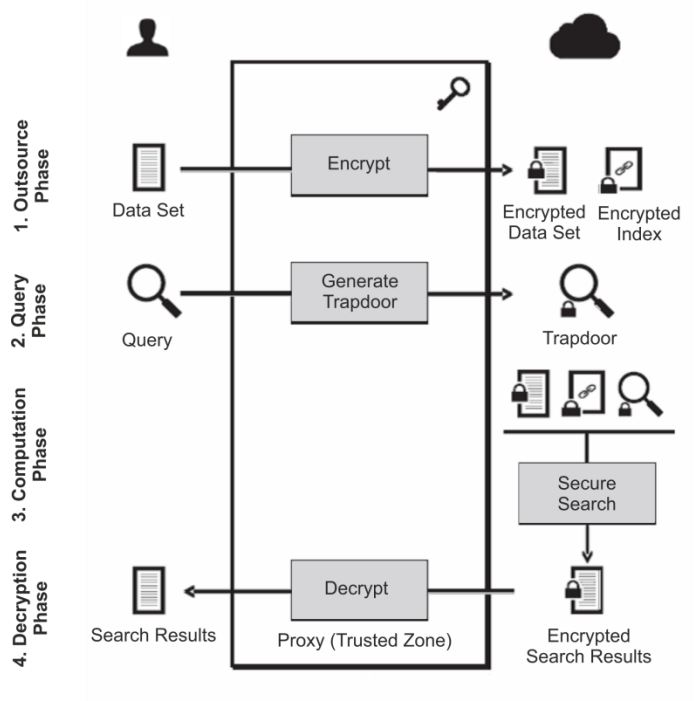

Fig.3.3 Encryption Workflow in Cloud Computing

The four key features expected to outsource encryption systems for cloud computing are:

- Scanning of encrypted data to retrieve segments of outsourced encrypted data to fulfill some query requirements. Since remote search functions play a key role in many commercial and industrial applications' business logic, solutions to securely search outsourced data can offer tremendous security benefits.
- Outsourced computer storage is a core component of cloud computing. Because of the high-performance gains and the ubiquity of information and functionality of cloud infrastructures, it is very easy for computers to carry out in the cloud in different applications. E-voting, financial systems and accounting, as well as measuring statistical metrics, are also important examples.
- Control of access to encrypted data which includes limiting data exposure to a specified collection of allowed recipients. Monitor access to data in a variety of cloud computing services, such as file-sharing or audit log systems, requires security policies. However, the cloud is vulnerable to certain privacy breaches by having an access management feature. They are also researching new approaches to handle outsourced data through encryption using user-centered access management.
- Outsourced data management access ensures that the customer can verify their outsourced data for credibility and location.

The encrypting schemes in cloud computing have been given in Figure 3.3 for the database server environment. Initially, as with all encryption systems, a setup algorithm is run locally to send a key to the client. The client then uses this key in an outsource stage (see Figure 3.3) to produce an encoded version of the original data set, consisting of a ciphertext compilation and an encoded index, which is outsourced to the server. The client may wish to collect the portion of the outsourced data set that matches a specific query during the query stage (stage 2). In this context, the client produces an encrypted, usually called a trapdoor and sends it to the server using the encryption scheme. If the trapdoor is released, the cloud will enter the processing process (stage 3) and merge this trapdoor with the encrypted index so that objects that match the database criteria are indexed within the encrypted data collection. After these coded responses are returned to the database, the database is decrypted (stage 4) and then the requested response is eventually obtained.

Advantages

- Authentication is valuable, as applicable to search situations, for those applications that require a protected search. Both the security assurances and approved queries of all these implementations vary.
- As the encryption gives the potential for all cryptographic techniques the most practical strategy for having a secure delegated scan.
- Here, the full security capabilities are accomplished and most outsourced data and all query/search information are secured.
- Also, it offers diverse data collection service solutions.

Disadvantages

- Encryption allows both the addition and multiplication of ciphertexts but permits a small number of operations in practice.
- A significant number of additions and a limited number of ciphertext products are accepted by most schemes. This limits functions that can be externalized to the cloud and thus restricts the appropriateness of encryption systems for such purposes.
- The overhead measurement methods are sadly too high for many implementations to be practical.

4. Results and Discussion

Here, we compare the types of security strategies discussed above in the given methodological criteria. The criteria required to validate the reliability of cloud computing approaches are supported operations, overhead at the local proxy, preservation of the accuracy of the original data, transparency and interoperability.

4.1 Supported Operations

All methods allow safe data to be produced that can be stored in the cloud. The variations are in the more activities supported. Although basic encryption stops the CSP from receiving encrypted data from working, the other two approaches are more reliable.

4.2 Overhead at the Local Proxy

The work done by the local proxy of the user differs according to the security process. To produce the results, the majority of the methods need different degrees of the number of records. The proxy would not need any help until the data set is created and submitted to the cloud.

4.3 Accuracy Preservation

Except for anonymization, the other approaches offer the consumer maximum precision. This is since both encryption and splitting are transformations that can be reversed. Since anonymization requires certain permanent modifications, a certain calculation of the anonymized data set which generate results which would not necessarily be obtained in the original data set.

4.4 Data Protection Level

Non-cryptographic approaches are given fewer secure data than encrypted approaches. Data replication generates datasets of original data when data anonymization processes outsource the entire data in a manner that is coarsened or skewed (according to a particular anonymization mechanism). The proxy security system is important in all cases to prevent threats of exposure.

4.5 Transparency

Anonymisation is fully transparent: no metadata or key information has to be stored on the anonymized data by a proxy and the CSP will work on it as if it were an original data. For the splitting of data, multi-cloud or multi-cloud

accounts are required, and the proxy must watch where each component was stored, while the fragments of the CSPs stored do not have to handle or do additional work. Thus separating the CSPs, but not the proxy, is clear. Encryption methods share the need for key proxy management that would hold all used keys. The proxy thus lacks transparency.

4.6 Interoperability

Anonymization between different networks (proxies or clouds) is conveniently interoperable. The interoperable mechanisms should not be aware of the anonymizing approaches used by the data (they will handle anonymized data as original). When splitting files, interoperable proxies must share the metadata that displays the location for each fragment. In terms of security, interoperability involves the exchange of the encryption keys between interoperable proxies. This is a simple limitation since if the key is corrupted during delivery, it can no longer protect the outsourced data.

Table 1: Comparison of Survey Methods

Method	Supported Operation	Overhead at the Local Proxy	Data Accuracy Preservation	Data Protection Level	Transparency	Interoperability
Anonymization	Storage, Queries on non-masked data, Any computation	Quasi-linear w.r.t data size during storage, Zero in all other operations	May be partial, Depends on the calculation requested and the masking method, Same for end-user and CSP	Data still in clear, but coarsed/distorted in an irreversible way	Transparent both for the proxy and CSP	Straight forwardly interoperable with an external system
Splitting	Storage, Updates, Queries, Computation on marginal, Joint computations require specific solutions	Constant for all operations	Full for end users, Full for CSP on the partial data it stores	Partial data in the original form, Could be broken via confabulation or compromised missed metadata on the splitting process	A multi-cloud or several cloud accounts are needed, Operations are transparent for CSP	Requires exchanging metadata on the splitting process
Encryption	Storage	Linear w.r.t data size on all operations	Full for end users, None for CSP	Fully encrypted data, Could be broken via compromised keys	Key management at the local proxy	Requires exchanging keys

5. Conclusion

Cloud data security is an important part of Cybersecurity and poses a significant challenge because it is focused largely on Internet access that leaves Cloud storage systems vulnerable to a spectrum of attacks and safety risks that can lead to mild and serious consequences. Although in cloud computing there are many safety problems, we addressed some of them and strategies for avoiding them in this article. They can be used for ensuring reliable connectivity and mitigating security issues. This study essentially discusses all issues, such as threats, lack of data and unauthenticated access to data, as well as the methods of elimination. Owing to the nature and complexity of cloud computing, standard cloud security solutions do not mesh up with their virtualized worlds. In this survey article, we reviewed some main approaches to offer cloud computing security. Also, we give options for the comparative study as well as potential countermeasures.

References

[1] Herman, M., Iorga, M., Salim, A. M., Jackson, R. H., Hurst, M. R., Leo, R., Sardinas, R. (2020). NIST Cloud Computing Forensic Science Challenges. doi:10.6028/nist.ir.8006

[2] Bhardwaj, A., & Goundar, S. (2020). Cloud Computing Security Services to Mitigate DDoS Attacks. Cloud Computing Security [Working Title]. doi:10.5772/intechopen.92683

[3] Gupta, B. B., & Badve, O. P. (2016). Taxonomy of DoS and DDoS attacks and desirable defense mechanism in a Cloud computing environment. Neural Computing and Applications, 28(12), 3655-3682. doi:10.1007/s00521-016-2317-5

[4] Yousif, M. (2017). The State of the Cloud. IEEE Cloud Computing, 4(1), 4-5. doi:10.1109/mcc.2017.4

[5] Liu, T., & Wu, G. (2018). Universal SaaS platform of internet of things for real-time monitoring. doi:10.1063/1.5033740

[6] Stevenson, D. M., & Pasek, J. (2015). Privacy Concern, Trust, and Desire for Content Personalization. SSRN Electronic Journal. doi:10.2139/ssrn.2587541

[7] Shynu, P. G., & Singh, K. J. (2016). A Comprehensive Survey and Analysis on Access Control Schemes in Cloud Environment. Cybernetics and Information Technologies, 16(1), 19-38. doi:10.1515/cait-2016-0002

[8] Zhang, R., Ma, H., & Lu, Y. (2017). Fine-grained access control system based on fully outsourced attribute-based encryption. Journal of Systems and Software, 125, 344-353. doi:10.1016/j.jss.2016.12.018

[9] Zhang, J., Chen, B., Zhao, Y., Cheng, X., & Hu, F. (2018). Data Security and Privacy-Preserving in Edge Computing Paradigm: Survey and Open Issues. IEEE Access, 6, 18209-18237. doi:10.1109/access.2018.2820162

[10] Salman, T., Zolanvari, M., Erbad, A., Jain, R., & Samaka, M. (2019). Security Services Using Blockchains: A State of the Art Survey. IEEE Communications Surveys & Tutorials, 21(1), 858-880. doi:10.1109/comst.2018.2863956