A Survey on Approaches of Data Confidentiality and Integrity Models in Cloud Computing Systems

1 Noha Alsulami, 2 Eman Alharbi, 3 Muhammad Mostafa Monowar

1, 2, 3 Information Technology Department, Faculty of Computing & Information Technology, King Abdulaziz University, Jeddah, KSA

naalsulami2@kau.edu.sa, 2 EAL-HARBI0001@stu.kau.edu.sa, 3 mmonowar@kau.edu.sa

ABSTRACT

Cloud computing is a technology that uses the Internet and central remote servers to access data and applications. Due to its public availability, the major challenge that faces cloud computing is how to secure, protect and process the data of the user. In this paper, we investigate different models and solutions for achieving the required level of confidentiality and integrity, according to recent studies. These models have been organized according to the security issues that we are handling. We found that there are many techniques that can be used to enhance integrity and confidentiality on the cloud, which include: encryption, separation, multilayered architectures, and anonymization. Encryption was found to be the most widely applied technique among the proposed models. There are also many models that combine more than one technique to gain a higher level of protection. It is worth mentioning that the security of the cloud computing environment requires further academic and industrial researches.

Keywords: Cloud computing, Confidentiality, Integrity, Encryption, Multilayered Architecture, Separation, Anonymization

1. INTRODUCTION

In the past, data and software had to be stored and processed on the same computer. But nowadays, cloud computing allows for a functional separation between the resources used and the user's computer, which makes everything more flexible and easier. The principle behind the cloud is that any computer connected to the Internet is connected to the same applications, computing power and files. Users can store and access personal files, such as pictures, music, videos and bookmarks, or do word processing on a remote server rather than physically carrying around a storage medium such as an MP3 player or a DVD. It became a solution for the rising storage and maintenance costs for IT enterprises and it enables customers to increase the capability and capacity of their machines on the go.

Cloud computing is built on established trends for driving the cost out of the delivery of services, while increasing the speed and agility with which services are deployed. Cloud computing incorporates on-demand deployment, virtualization, Internet delivery of services, and open source software [1]. From one perspective, cloud computing is nothing new, because it uses concepts, approaches and best practices that have already been established. From another perspective, everything about it is new, because cloud computing changes how we invent, deploy, scale, develop, maintain, update and pay for applications and the infrastructure on which they run.

Many companies, such as Amazon, Google and Microsoft, are accelerating their pace in developing cloud computing systems and enhancing their service provision to a larger proportion of users. The successes of these companies serve as great examples and have encouraged a number of other companies to step into the cloud, such as Media Temple, Mosso, Joyent and Flexicale [2]. In 2007, IBM and Google announced a collaboration in cloud computing [1]. Besides the web email, the Amazon Elastic Compute Cloud (EC2), Google App Engine and Sales force’s CRM largely represent a promising conceptual foundation for cloud services. There are characteristics of cloud computing that distinguish it from other systems that depend on the Internet: [3]

a. Resource sharing: Resources in cloud computing can be shared among different users and the same resources can be used in the same network level, depending on the business.

b. Scalability: In cloud computing, storage and the number of users have the ability to expand.

c. Flexibility: Cloud computing provides users with flexibility, based on the decreasing or increasing amount of resources used.

d. Self-identifying resources: The appropriate capacity for subscribers in cloud computing is based on software, storage and processing.

e. Pay as you use: The cost in cloud computing is calculated depending on its use.

However, data security over the cloud is a major concern and it is considered as biggest challenge, since its resources are centralized and the levels of security provided by each cloud provider are not the same. Therefore it is necessary to implement best practises, in order to provide the best potential level of security for users in the cloud.

Recently, various models have been proposed to introduce a solution to the security problem in the cloud. The main concern of these models has been dealing with data integrity and confidentiality in the cloud. “Data integrity” refers to the protection of information from modification by unauthorized parties. “Data confidentiality” means protecting the information from disclosure to unauthorized parties.

In this paper, we have investigated many different models and technologies that deal with the security issues in cloud computing, especially those of
integrity and confidentiality. We made a comparison between these models, and found that the use of encryption plays a major role in ensuring data integrity and confidentiality. However, another approach was shown to assure information and data confidentiality, which involved the use of multilayered architecture with agents, file permissions and access control lists to limit access to the sensitive information.

The rest of this paper is arranged as follows: Section 2 presents an overview of cloud computing; Section 3 introduces elements and models for security in cloud computing; Section 4 conducts a comparative study of the existing models; Section 5 presents the research challenges and future directions; and Section 6 concludes the paper.

2. OVERVIEW OF CLOUD COMPUTING

In this section we will present an overview of the architecture and security requirements of cloud computing.

2.1 The Architecture of the Cloud

In general, geographical distribution works to promote low cost, scalability, flexibility, homogeneity, advanced security technology and service-orientation. Computing architecture is classified based on the type of cloud service.

2.1.1 Cloud Service Delivery Models

There are three main categories of model, as shown in Figure 1 [4]:

A. Software-as-a-Service (SaaS)

SaaS is a set of IT applications run in the cloud and is used to deliver applications through a browser to thousands of customers, using a multiuser architecture. SaaS focuses on the end user/client as opposed to managed services. Basically, a managed service is an application that is accessible to an organization’s IT infrastructure, rather than to end users. Examples include Gmail from Google, and Microsoft Office Live.

B. Platform-as-a-Service (PaaS)

PaaS is closely related to SaaS, but delivers a platform on which to work, rather than an application to work with. It provides the capability to build or deploy applications on top of an infrastructure layer. PaaS delivers development environments to analysts, programmers and software engineers as a service. A provider of cloud computing offers multiple application components that align with specific development models and programming tools.

C. Infrastructure-as-a-Service (IaaS)

IaaS is a model in which an organization outsources the equipment used to support operations, including hardware, servers and networking components. IaaS is similar to the managed services offered in the Internet era (e.g. storage service providers, hosting services, etc.).

2.1.2 Cloud Deployment Models

There are four deployment models, according to the models of cloud service delivery: [3]

a. Public cloud, where the cloud infrastructure is available to a group of organizations;

b. Private cloud, where the cloud infrastructure is available to a single organization;

c. Community cloud, where the cloud infrastructure shares a set of characteristics in the cloud, and is shared between many organizations; and

d. Hybrid cloud, where then cloud infrastructure consists of two clouds or more, as mentioned above.

2.2 Security Requirements in Cloud Computing

There are two categories of security related to cloud computing: [5]

2.2.1 Cloud Computing Security

The integrity, confidentiality and availability of a system are guaranteed by securing the infrastructure of a cloud computing system.

2.2.2 Cloud computing for Security

Security solutions that are developed and delivered for IT systems, using technologies of cloud computing.

One important consideration in cloud computing concerns service providers, and the ways in which users can protect their confidential data from being disclosed to them. Users should ensure that their service provider does not know where their confidential data is located in the cloud, and that they can’t understand the meaning of this data, as well as having no privilege to access or collect it. The security requirements in cloud computing include: [6]

a. Confidentiality: assuring that only authorized users or systems have access to the data;
b. Integrity: assuring the non-modification or manipulation of the data stored in database or in its transmission state, except by trusted persons or processes;

c. Availability: assuring the data’s existence for the authorized users and systems;

d. Authenticity: assuring that query transactions, contracts and communications are genuine, and that involved entities (users/system) are known and verified;

e. Accountability: assuring that all valuable assets are protected from illegal access (this is also called access control); and

f. Risk management: which includes a number of activities to track and identify data security vulnerabilities, as well as the control measures set up to avoid further risk.

3. ELEMENTS AND MODELS FOR SECURITY IN CLOUD COMPUTING

Considering the diverse and inherent complexity of cloud computing environment, conventional security methods are not always applicable on the cloud computing. Instead, new policies should be designed specifically for cloud security. There are certain techniques or elements that have a particular effect upon the cloud security issues related to data integrity and confidentiality. It should be noted that the implementation or mere existence of these techniques in any cloud system are not mutually exclusive. They can coexist within one system to enhance the security level. Figure 2 shows these techniques.

![Diagram](image)

**Fig 2:** Techniques contributing to cloud security

In this section we will investigate various models that have been presented in recent studies, which can enhance data integrity, data confidentiality or both of these, in the cloud. Most of these models rely on the use of encryption techniques, while others use a combination of more than one technique, in order to offer high security on the introduced service.

Encryption works by translating the original/plain text into cipher text, which means encoding the message so that unauthorized users cannot read it.

There are various encryption algorithms, some of which use a single key that is shared between the parties, while others are based on two keys: a public and a private key. The latter is the type used in the proposed scheme; it provides a strong, secure environment, since the private key generation depends on the use of a module which produces two large prime numbers. In this way, the key should not be compromised, as an attacker would need a long time to complete the factorization operation necessary to find out this key. The encryption process which uses such keys is called an “asymmetric cryptosystem”. In the following, we will discuss various encryption techniques that are used, such as RSA, homomorphic, and others. Some of these techniques use the hash function with the encryption, to offer more security around the data.

Layered architecture can be used to allocate different responsibilities to each layer and to add more protection to the data. Some of the proposed models were applied to the agents in their layers, which were used to perform special tasks.

The literature describes a different solution to achieving security in cloud computing, by applying a separation model. This is a cloud services deployment model that is designed to ease the security concerns of users by separating data processing from data storage, using different service providers.

Moreover, anonymization is another effective technique that protects cloud computing systems against various attacks. This is the alteration of the original data in such way as to prevent the exposure of this data. It also limits information loss.

3.1 Data Integrity Models

There are many models that have been introduced to enhance the data integrity of cloud computing and to offer users the ability to check the correctness of their data, to ensure that no modification has occurred.

In [7], a scheme was proposed for checking data integrity. The main feature of this scheme is that it does not involve the encryption of the whole data; it only aims to encrypt a few bits of data per data block. The integrity of the data is measured by verifying the meta data, which is preprocessed by the verifier before uploading the file, then appended to the file itself. This was compared with the scheme called Proof of irretrievability for large files using "sentinels", and it was found that this scheme surpasses the other schemes. It reduces the network bandwidth consumption, and the computational and storage overheads of the client, as well as minimizing the computational overhead of the cloud storage server, which proves its advantageousness to thin clients, like mobile phones. The limitation of this scheme is that it can be applied only to static storage of data, and it can check the integrity without preventing unauthorized modifications.

Another scheme for data integrity checking was proposed by Jianhong et al. [8]. This scheme utilizes the RSA security scheme by combining it with a digital
signature. Here, data is divided into blocks of variable size and these block tags are authenticated for integrity checks, instead of the original data blocks. The obvious advantage to this scheme is that the client does not need to keep the copy data on their side, so it certainly relieves the storage burden of the client. However, this scheme cannot realize the dynamics data for a remote data integrity check.

A fine-grained data integrity check, based on combination and error-correcting code, was introduced by Yang et al. [9]. This method can compress effectively detection value for data integrity, to reduce storage and improve detection efficiency of multi-error data objects. \( N \) data objects are combined, and a check matrix is obtained. Then, each combination is calculated using a one-way hash function, and the results are stored in a hash set. This study provides a good theoretical support for the security application of power cloud computing.

A layered architecture named “Cloud Zone” [10] represents an integrity layered architecture of a typical cloud. It is based on multi-agent architecture, which increases the reliability and efficiency of this architecture, and consists of two main layers. The first is the cloud resources layer, where storage servers and cloud application servers are located. The second is the multi-agent system’s architecture layer, which in turn consists of two agents: the Cloud Service Provider Agent (CSPA), which provides a graphical interface to the cloud user and facilitates the access to the services; and the Cloud Data Integrity Backup Agent (CDIBA), which enables the cloud user to rebuild the original data by using SQL programming. The main responsibility of this agent is backing up the cloud data regularly from “Cloud Zone”, and sending security reports and/or alarms to the CPSA when human errors occur, as well as errors occurring when cloud data is transmitted from one computer to another, from software bugs or viruses, and from hardware malfunctions such as disk crashes. This architecture has been analyzed through all the system development life cycle phases, and it has been tested and validated. However, this architecture works only with the MS SQL database, and it does not support any other type of files, nor does it support component-based backup.

Recently, Map Reduce has become the most successful technique used for processing parallel data-intensive applications. It has been popularized by many large companies. It is capable of representing several types of application, that can be reduced to two functions: mapping and reducing. This technique is suitable for use in grids, clusters and cloud computing environments. However, the deployment of Map Reduce in cloud computing systems presents several challenges when considering cloud computing as an open public system. Since users of Map Reduce in the public cloud have no control over the workers where their data is submitted, given that these workers may not always be trusted, and due to the use of public networks for communication and data transfer, Map Reduce faces a wide variety of security threats and security risks, especially for data integrity. To address this problem, a verification mechanism based on a replication-based voting method and a reputation-based trust management system has been introduced in [11]. The basic idea of this mechanism is that each task is replicated and executed by multiple workers, and the results are collected and grouped according to the value of the results. However, the mechanism also takes into account the overall computing behavior of each worker, expressed by a worker reputation value as the trustworthiness of the worker. In this mechanism, the replication and execution of a task is repeated until one result group reaches a group reputation higher than the weight threshold, and subsequently the result value returned by the workers of this result group is accepted as a correct result. In addition, the reputation of these workers is updated each time the returned result is accepted as correct. This mechanism can efficiently detect both collusive and non-collusive malicious workers, and guarantee high computational accuracy with an acceptable overhead. The limitation of this mechanism is that it has been used to analyze one type of malicious worker only, which is the collusive type.

3.2 Data Confidentiality Models

There are various proposed models that can ensure the confidentiality of data in cloud computing. Encryption is the most common technique used to ensure confidentiality in the cloud environment. One of the models proposed to ensure the confidentiality of data transferred to cloud storage is called Cloud Storage Encryption (CSE) Architecture. It uses the access policies of the encryption and decryption process, and searchable encryption. Encryption techniques in this architecture are able to provide a high level of data protection during data transfer to cloud storage. However, the performance of this architecture is not tested and integrity is not guaranteed in this architecture [3].

Another approach, called “information-centric”, is presented by Valentine et al. [12]. It refers to protecting the content of data itself to extend control of data in the cloud, and to maintain confidentiality and trust in distributed data storage by using the encryption of data to secure it. This approach solves various security problems by building trust. However, there has been no implementation of the proposed approach. In [13], another system model, called “attribute-based encryption”, is proposed. It divides and stores each data file into header and body in the cloud. The body is formed by using a Data Encryption Key (DEK) to encrypt the file before storing it in the cloud, while the header is formed by using attributes to encrypt the DEK. Then the header is sent to a trusted authority called a “privilege manager group”, and the body is sent to the cloud service provider. This model is more secure against collusion attackers, who illegally read data by colluding with the cloud server. However, this model requires more effort in order to handle and communicate the separated parts, and integrity is not taken into consideration in this architecture.
Although encryption is the most common technique used to ensure confidentiality in the cloud environment, it is not enough for data protection. Obfuscrypt is a newly proposed confidentiality technique to protect data in the cloud storage from various attackers. It is an integration of obfuscation and encryption. Obfuscation is a process that uses programming techniques or implements a particular mathematical function to disguise unauthorized users. Obfuscation is applied to numeric data, while encryption is applied to symbols, alphanumeric and alphabetic data. This technique requires the entire user's data to be obfuscated and encrypted before it is sent to cloud storage. Obfuscrypt is the newest technique offering better security to stored data in the cloud environment than other techniques which use obfuscation or encryption alone. However, it does not involve any integrity verification mechanisms [14].

The literature presents different solutions for achieving confidentiality in cloud computing, by applying a separation model. One of these solutions is the Redundant Array of Independent Net-storages (RAIN) that is proposed in [15]. It is an approach that splits the information into segments and distributes these segments amongst multiple storage providers. Hence, disclosing only some of these segments will result in no meaningful information to others. This approach ensures the confidentiality of data stored on clouds, without depending on heavy cryptographic operations and without having to trust each individual provider in cloud computing. The limitation of this approach is that no experiments or practical prototypes have been done to validate it yet.

Existing techniques of data fragmentation are not specially designed for data security. To achieve confidentiality of data outsourced to cloud environments, one proposed method uses a fragmentation technique while applying minimal encryption. It uses normalization of relational databases, and categorization of tables based on user requirements, related to performance and availability, and exported to XML as fragments. Then, the lowest necessary number of Cloud Service Providers (CSPs) is required to store all fragments that must remain unlinked in separate locations, which occurs after defining the fragments and assigning the appropriate confidentiality levels. However, this method does not guarantee data integrity [6].

Service providers represent the most serious threat to data confidentiality in the cloud environment. Anonymization is an effective technique for confidentiality protection against service providers in cloud computing systems. There are many solutions that employ anonymization to ensure confidentiality. One of these solutions is called a High-Performance Anonymization Engine (HPAE). It is a practical approach that enables organizations to implement their own controls to protect sensitive information from service providers when data is processed in the cloud environment. The architecture of HPAE contains different components, such as techniques of anonymization, more efficient libraries/data structures, and input/output types to handle large volumes of data that offer high-performance and fast processing. The performance of this approach is assessed by implementing a prototype that shows HPAE's architecture to be effective practical solution, because it can handle large volumes of data. However, this prototype has not yet been completed at the time of writing, and it does not contain all the components of HPAE. Additionally, integrity is not guaranteed by this architecture [16].

Another approach that implements anonymization is proposed by Yau et al. [17] to protect users' confidential data from cloud service providers and ensure that this data cannot be collected by service providers when it is processed and stored by a cloud system. This approach depends on three features: 1) separate infrastructure and software service providers in cloud computing enable different service providers to manage them separately; 2) the information about data owners is hidden, by employing data anonymization; and 3) data obfuscation is used to process user data without exposing confidential information about them to infrastructure service providers in the infrastructure cloud. However, this approach does not always work properly, due to the fact that the software that manages the platform or infrastructure service is offered by the provider. Integrity is also not guaranteed in this approach.

### 3.3 Hybrid Models

Some of the proposed models tend to use a combination of different approaches to offer data integrity and confidentiality in the cloud.

A new verification protocol that can accommodate dynamic data files and authenticate remote server storage data has been constructed in [18]. This solution is used to provide confidentiality and integrity to the users of cloud services. The design and development of this protocol is mainly based on the employment of an RSA public and private key encryption system, based on a hash function. It supports the dynamic outsourcing of information, making it a more realistic application for cloud computing. Confidentiality is guaranteed by applying the encryption process before storing the file on the server, ensuring that the file will not be intercepted by an unauthorized person attempting to obtain the file content. In the verification phase, integrity is guaranteed by calculating the value of the cipher text stored in the server with the cipher text stored on the client side. If the values are equal, it means that the server has the correct storage cipher text file. However, this protocol slows down the entire computing speed, due to the RSA-based cryptography.

Nirmala et al. [19] employ a user authenticator scheme to handle dynamic operations on the data. This scheme combines an encryption mechanism with a data integrity check mechanism, in order to cover the data
The data owner is responsible for generating the public and private key needed to encrypt the data. First, they divide the data files into blocks and encrypt each block using their private key to ensure confidentiality. Then, they compute the hash value for each block to generate the digital signature. After that, the digital signature will be added to each data block. The storage server will contain the encrypted data blocks with its signatures. In addition, the data owner has a copy of the hash codes for each block, in order to verify the integrity by requesting the hash code stored at the server side, and then comparing it to the one in their database. Moreover, the data owner generates the access codes and policies when another user wants to use their data, and sends it to the server to verify it. This scheme seems not to be efficient for either the data owner or the cloud server. It creates a burden for the data owner in generating the hash codes and the access codes and policies, and it makes the server busy with hashing update access code requests each time data is changed or accessed by another user.

The integer-based homomorphic encryption scheme, which was introduced by Ramaiah et al. [20], has been enhanced and applied to a cloud resident data integrity auditing protocol in [21]. The homomorphic encryption scheme consists of four algorithms: KeyGen, Encrypt, Decrypt and Evaluate. The new scheme was conducted as an attempt to combine data integrity and confidentiality in new ways, to provide an adjacent solution to the cloud security challenge. In actuality, this protocol provides the integrity guarantee, while data retrieval may or may not be guaranteed. The practicality of the resulting scheme is demonstrated through an exhaustive experiment, while the security of the scheme has been proven with thorough cryptanalysis. Experimental results substantiate the performance of the schemes. Research is also actively underway to develop cryptographic trust mechanisms, such as the verification of the computations performed on the cloud resident data (called Verifiable Computation (VC)). However, it does not support public auditability and dynamic data change.

### 3.4 Models Concerned With Other Security Issues

There are other security issues encountered in cloud computing, in addition to data integrity and confidentiality, such as privacy, data leakage and data availability.

A 3D framework was introduced in [22] to provide a solution to data leakage in cloud computing. It is based on assigning a priority to each data added by the client, according to his/her categorization. Data is categorized according to “CIA” (confidentiality (C), integrity (I), and availability (A)), in which the client assigns a value C, I and A. The value of C is the level of secrecy, the value of I is the provided assurance of accuracy and reliability of information, and the value of A is based on how frequently it is accessible to other authenticated users. Then, the data that has the highest rating is considered to be critical. The allocation of data on the basis of the rated priority is performed in a 3D protection ring, in which the internal protection ring is very critical and requires more security techniques to ensure confidentiality. Later, the execution of a user's request will be performed according to the protection ring which includes the data. If the protection ring is “1” then a strong authentication is required. The authentication process will be validated by the authority organization which stores all the registered users; if the user is authenticated then it will redirect him/her to the cloud provider to access resources. This technique provides a new way of authenticating within a 3D approach. It guarantees the availability of data by overcoming many existing problems, such as denial of services and data leakage. In addition, it also provides more flexibility and capability to meet the new demands of today’s complex and various networks. The proposed framework complies with the data protection and privacy legislation of most nations, including India’s Section 43A of the Information Technology Act 2008, the USA’s Health Insurance Portability and Accountability Act of 1996 (HIPAA), and the UK’s Data Protection Act 1998. However, this framework is still under observation, since it has not yet been applied to a real world cloud provider or tested by a simulated project.

Another of these models, introduced by Zhang et al. [23], works by dividing cloud computing protection into three layers: the first layer is authentication protection, the second is file encryption and privacy protection, and the last layer is file fast recovery. Agents have been introduced for each of these protection layers. Each agent has its duties. The authentication agent, which lays on the authentication protection layer, is responsible for authentication, managing user permissions and issuing digital certificates. The encryption agent, which lays on the file encryption and privacy protection layer, selects the encryption algorithm and protects the key. Finally, the privacy agent, which lays on the file fast recovery layer, protects users’ privacy. This proposed model can prevent any attack because it operates as three associated defense walls; if an unauthorized user were to trick the authentication agent, it would come to the encryption agent, who would take care of the key. If the key were stolen, the privacy agent would protect the private data by checking whether they were the proper user, according to its knowledge. However, the concept of privacy here is from the perspective of keeping the information private from the attacker, not from the service provider themselves, while they still have control over that agent. Moreover, all these processes can cause delays to user requests and affect the speed of processing in the cloud. Additionally, this model requires further verifications, which are mentioned by the authors of [23] as possible future work.

Data privacy is a special aspect of data confidentiality that means the data owned by an individual will never be disclosed to unauthorized users. Encryption of a user's data is the most effective and simplest way to ensure data privacy [6]. This is achieved by applying searchable encryption techniques that take a keyword that
is sent encrypted to the cloud. The cloud will then return the result, while the actual keyword for the search will not be known by the cloud. The application of fuzzy keyword search schemes to encrypted cloud data has been proposed, in order to achieve privacy in cloud computing. This involves constructing fuzzy keyword sets by combining edit distance with wildcard-based techniques, to address the problems of format inconsistency and minor typos, and to reduce representation and storage overheads. However, it does not handle large volumes of data in real time, and integrity is not guaranteed by this scheme [24].

A distributed storage system is considered a big problem for data security in cloud data storage. To enhance the security of data storage in the cloud, an effective and flexible distributed scheme called “Defense in Depth” has been proposed, which is a layered approach to providing a comprehensive framework based on security policy. The integration of storage correctness insurance and data error localization, which is one of the main conditions for eliminate errors in storage systems, is achieved in this scheme. As result, storage in a cloud environment will be more resilient to future attacks. However, this scheme does not ensure the reliability of data retrieval from cloud storage, due to the fact that data is split into many blocks in order to be stored [25].

4. COMPARATIVE STUDY OF THE EXISTING MODELS

As we have seen in the previous section, while security is the major concern of the cloud computing environment, there are many models that have been introduced to provide more secure and reliable services to users of the cloud. The advances in security models are being developed to guarantee the best acquired solution for all security issues, in addition to enhancing the efficiency of network and storage requirements. Since users are looking for the model that provides the best practices for them, the most efficient and secure model is one which can handle more than one security issue, without burdening user storage or causing a connection delay. From this perspective, there is no one model which fulfills all these requirements.

Table 1 presents a list of proposed models, with their security techniques, strengths and limitations. We can say that the most secure model (from the perspective of protection) is the one proposed in [23], since it meets the main security challenges, which are confidentiality, integrity and privacy. On the other hand, the most efficient model (from the perspective of storage and connection) is the one proposed in [14].

<table>
<thead>
<tr>
<th>Literature No.</th>
<th>Techniques employed</th>
<th>Enhanced security issues</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>[7]</td>
<td>Encryption of a few bits per data block</td>
<td>Checking the integrity of the uploaded data in a light approach</td>
<td>Reduces the computation, network bandwidth and storage overheads</td>
<td>- Does not prevent changes in data - Does not support dynamic data change</td>
</tr>
<tr>
<td>[8]</td>
<td>Encryption and digital signature</td>
<td>Ensuring integrity and confidentiality</td>
<td>Uses an RSA scheme</td>
<td>- Cannot realize dynamic data change</td>
</tr>
<tr>
<td>[9]</td>
<td>Hash function and error correcting code</td>
<td>Checking integrity and correcting errors</td>
<td>Supports the security application of power cloud computing</td>
<td>- Does not prevent changes in data</td>
</tr>
<tr>
<td>[10]</td>
<td>Multilayered architecture and agents</td>
<td>Ensuring integrity</td>
<td>Offers backup for the data</td>
<td>- Support special file types only</td>
</tr>
<tr>
<td>[11]</td>
<td>Replication-based voting method and reputation-based trust management system</td>
<td>Checking data integrity and preventing unauthorized modifications</td>
<td>Efficiently detects collusive malicious workers and guarantees high computation accuracy with an acceptable overhead</td>
<td>- Detects for collusive malicious workers only</td>
</tr>
<tr>
<td>[3]</td>
<td>Searchable encryption</td>
<td>Ensuring confidentiality of data transferred to cloud storage by using searchable encryption</td>
<td>Provides a high level of data stored in cloud storage</td>
<td>- Not tested - Cannot ensure integrity</td>
</tr>
<tr>
<td>[12]</td>
<td>Encryption</td>
<td>Maintaining confidentiality and trust in distributed data storage by proposing an information-centric approach that uses encryption to secure data</td>
<td>Solves the main security problems in distributed data storage by building trust</td>
<td>Has not been implemented</td>
</tr>
<tr>
<td>[13]</td>
<td>Encryption</td>
<td>Ensuring confidentiality by using attribute-based encryption</td>
<td>Secured against collusion attack</td>
<td>Requires more effort to handle and communicate separated parts</td>
</tr>
</tbody>
</table>

Table 1: The proposed solutions and their limitations
<table>
<thead>
<tr>
<th>Reference</th>
<th>Scheme</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>[14]</td>
<td>Encryption and obfuscation</td>
<td>Ensuring confidentiality to protect data in storage, by integrating obfuscation and encryption</td>
<td>Offers better security in the cloud and outperforms existing techniques that use obfuscation or encryption alone</td>
<td>Does not involve any integrity verification mechanism.</td>
</tr>
<tr>
<td>[15]</td>
<td>Separation model</td>
<td>Ensuring confidentiality of stored data in the cloud by using a RAIN approach</td>
<td>Achieved without depending on heavy cryptographic operations</td>
<td>Has not yet been validated by any experiments or practical prototypes</td>
</tr>
<tr>
<td>[6]</td>
<td>Encryption</td>
<td>Ensuring confidentiality of data outsourced to cloud computing by using fragmentation while applying minimal encryption</td>
<td>- Requires a low number of cloud service providers to store fragmentation - Minimizes the computations’ overhead by applying minimal encryption</td>
<td>Does not guarantee data integrity</td>
</tr>
<tr>
<td>[16]</td>
<td>Anonymization</td>
<td>Ensuring confidentiality by protecting user data from service providers using HPAE architecture</td>
<td>Provides an effective practical solution since it can handle large volumes of data</td>
<td>- Prototype not yet completed yet - Cannot ensure integrity</td>
</tr>
<tr>
<td>[17]</td>
<td>Anonymization and obfuscation</td>
<td>Ensuring confidentiality by a proposed approach that uses anonymization to hide user data</td>
<td>Prevents service providers from collecting user data</td>
<td>- Does not always protect user data from the cloud server - Cannot ensure integrity</td>
</tr>
<tr>
<td>[18]</td>
<td>Encryption and hash function</td>
<td>Ensuring integrity and confidentiality</td>
<td>Uses an RSA scheme</td>
<td>Slows down the system, creating a storage burden on client and server sides</td>
</tr>
<tr>
<td>[19]</td>
<td>Encryption, hash function and digital signature</td>
<td>Ensuring integrity and confidentiality</td>
<td>Applies various techniques to protect the data</td>
<td>Creates a data storage burden on the server and client side</td>
</tr>
<tr>
<td>[20]</td>
<td>Encryption</td>
<td>Ensuring integrity and confidentiality</td>
<td>Uses a homomorphic scheme.</td>
<td>Does not guarantee data retrieval</td>
</tr>
<tr>
<td>[22]</td>
<td>Uses a 3D ring, each weighted according to CIA</td>
<td>Preventing data leakage and Ensuring data privacy, confidentiality and availability</td>
<td>Ensures data availability and flexibility</td>
<td>Does not check the integrity of data.</td>
</tr>
<tr>
<td>[23]</td>
<td>Multilayered architecture, agents, encryption and digital signature</td>
<td>Ensuring privacy, integrity and confidentiality</td>
<td>- Prevents any attack, thanks to its three associated defense walls - Offers data recovery</td>
<td>- Does not guarantee privacy from the service provider - Causes delays in the network - Need more verification and analysis</td>
</tr>
<tr>
<td>[24]</td>
<td>Encryption</td>
<td>Ensuring privacy by applying a fuzzy keyword search scheme to encrypted cloud data</td>
<td>Reduces representation and storage overheads</td>
<td>Does not handle large volumes of data in real time - Cannot ensure integrity</td>
</tr>
<tr>
<td>[25]</td>
<td>Multilayered architecture</td>
<td>Enhancing the security of data storage by using an effective scheme called “Defense in Depth”</td>
<td>More secure against various attacks</td>
<td>Does not guarantee reliability of retrieved stored data</td>
</tr>
</tbody>
</table>

### 5. RESEARCH CHALLENGES AND FUTURE DIRECTIONS

Having obtained a deeper insight into the introduced schemes and models, we found that most of them pose a specific challenge that can be considered as obstacle, preventing them from being the best solution. The main challenges can be considered in the following points:

- Securing the user's data without burdening the storage on the cloud server or the user’s computer;
- Developing a model that offers data confidentiality, integrity and privacy;
- Using schemes which do not causing a delay in the network; and
• Handling various types of files and supporting dynamic changes in the data.

The future direction of cloud computing should face all these challenges and try to implement the best practice model, which offers the strongest security level.

6. CONCLUSION
The major challenges faced by the cloud computing community is how to involve methods of securing and protecting user data and data processes, given that these are publicly available. In this paper, we have investigated the different approaches that are outlined in recent studies for ensuring the integrity and confidentiality of users' data in cloud computing. We have analyzed them according to the enhanced security issues, wherein they have been organized into models that deal with data integrity, models that deal with data confidentiality, models that deal with both of these, and models that also deal with additional issues (primarily with data privacy and availability, and preventing data leakage). There are many techniques used by developers to guarantee security over the cloud, which include: encryption, separation, multilayered architectures and anonymization. Encryption technique were found to be the most widely applied in the proposed models. Finally, we have also presented a brief comparison between these models, according to their security issue solutions, strengths and limitations.

ACKNOWLEDGEMENT

We would like to thank Allah who helped us to achieve this paper. After that, we would like to thank Lecturer Hind Almisbah for her valuable comments and suggestions that helped to improve this research.

REFERENCES


AUTHOR PROFILES

Noha Ateeq Alsulami is a master student in IT Department at King Abdulaziz University, interested in agile development, cloud computing and information security.

Eman Talal Alharbi is undergraduate master student, information technology department, faculty of computer science in king Abdulaziz University. Her interests are research in Network Security Protocols, Cryptography and Cryptanalysis.

Dr. Muhammad Mostafa Monowar holds a Doctorate degree in Computer Engineering from Kyung Hee University, Yongin, Gyeonggi, South Korea. The research area of interest are: Wireless networks, especially ad hoc, sensor and mesh networks, including routing protocols, MAC mechanisms, IP and transport layer issues, and cross-layer design and QoS provisioning..