International Journal of Research in Advanced Engineering and Technology
www.allengineeringjournal.in
ISSN: 2455-0876
Received: 06-10-2023, Accepted: 22-10-2023, Published: 07-11-2023
Volume 9, Issue 4, 2023, Page No. 14-18

Analysis of hybrid data integrity protection techniques

Okozor Nkeiruka Petrolina, Okezie Christian Chikodili
[1] Department of Electronic and Computer Engineering, Nnamdi Azikiwe University, Awka, Nigeria

Abstract
Hybrid techniques are the act of using more than one authentication method to prevent unauthorized persons from having access to stored data. In this paper the analysis of Hybrid Data Integrity Protection Techniques was carried out. This technique comprises of data validation, digital signature and data encryption. The pros and cons of each of these methods were analyzed and summarized. Observation of the results proved that no matter the advantages associated with each method, there is still some disadvantages associated with each technique. This should serve as a guide in using any of the methods.

Keywords: Hybrid data integrity, data validation, digital signature, data encryption

Introduction

Security in today's world is one of the important challenges that people are facing all over the world in every aspect of live. Similarly security in electronic world has a great significance. The area of substantial interest in this paper is hybrid data integrity protection techniques, as data is a valuable asset for any organizationwhether social, governmental, education etc. for decision making. Hybrid techniques which are using more than one authentication method to protect the confidentiality of sensitive data stored in a repository or anywhere. There is need to use a hybrid techniques to avoid unauthorized persons from having access to crucial information, here we are looking at analysis of this hybrid techniques used which includes data validation, digital signature and data encryption techniques. Hybrid technique is more secure as compared to previous techniques. It is a combination of more than one authentication method to secure data so that it will not lose its integrity. According to Iqra el al (2012) database security demands permitting or prohibiting user actions on the database and object inside it. Organizations that run a successful business demand the confidentiality and integrity of their data. Data integrity protection is important as modified data or information is of no use to any organization. Often, it happens that important information or data is leaked out or misused not because of defective access control but because of improper information flow. When polices for information flow are not properly defined then the system data is less protected. For this reasons it is important to know the pros and cons of each method and know whether to use or not,

1. Data Validation

Data validation means checking the accuracy and quality of source data before using it. It can also say that it is a process of checking data that meets requirements by comparing it to a set of rules that have already been set up or defined. This procedure entails performing a series of checks known as check routines. The goal of data validation is to create data that is consistent, accurate and complete so as to prevent data loss and errors during a move. Different types of validation can be performed depending on destination constraints or objectives. Data validation is a form of data cleansing, is a general term and can be performed on any type of data. These are needed because it is easier to try and prevent users from entering garbage than attempting to fix mistakes later. Where data integrity is an objective it is imperative that the data be validated before acceptance for storage. If the original data lacks integrity ab - initio, nothing can make it have integrity after storage. With computers, it is garbage in, garbage out. It is important that the data supplied must match the field type defined in the database structure.

2. Digital Signature

A digital signature is an encrypted file that travels with the electronic document that needs to be signed and returns with it after the transaction has been completed. Digital signatures are among the most important components of an e-signature program, and they can drive security, legal validity, and records management efficiency when using an electronic signing method. Electronic signatures are making life easier for records management leaders, human resources employees, and managers in a variety of industries. However, digital signature solutions have not become pervasive everywhere, but the foundational technology is in place to drive innovation in even some of the strictest business sectors. Digital signature is a digital code which is authenticated by public key inscribed which is attached to an electronic transmitted document to verify content and sender's identity. The digital signature is the technique for approving legally the authenticity or virtue of message or documents. Digital Signature is equally valid as well as legal as the self attested or handwritten signature.

3. Encryption

The word encryption is commonly used to describe techniques that are intended to protect information from unauthorized access. Various methods have been used to transform readable content known as plaintext into a non-readable form known as cipher, from ancient times. The complexity of encryption methods has been successively

increasing ever since. Today's cryptographic methods protect data with the aid of computers using very long key lengths, complex mathematical tasks, and sophisticated key exchange procedures. As a vital element of information security, encryption helps to secure electronic communication between people and devices from spying and manipulation. Encryption is a method commonly used in IT to protect data from unauthorized access. To do this, algorithms turn the data from a readable into a non-readable format. The information is now only accessible to those who possess the keys needed to decrypt the data. There are two forms of encryption, symmetrical and asymmetric. But there are several types of encryption as well as different algorithms to aid the encryption process. The benefits of encryption include increased security and privacy, but a disadvantage is that this system requires detailed planning and maintenance. Although it seems like common sense to use data encryption in business and other entities for security, many organizations are opposed to encrypting data because of some of the obstacles involved with doing so.

Literature Review

PriyankaDeore and TusharChaudhari (2017) worked on Hybrid Encryption for Database Security. In this technique first, they encrypt plain text database by using symmetric key means secret key and stores that encrypted database in cipher text database. Then secret key which was used to encrypt database need to make secure by using public key, so asymmetric encryption technique was used. This complete encrypted block which contains encrypted database and encrypted secret key is send to the receiver site. Once receiver receives this encrypted block first encrypted secret key is decrypted by private key. Now decrypted secret key is used to decrypt encrypted database.

KatanoshMorovat (2015) worked on data integrity verification in cloud computing. The research work outlines how data owners determine which data should be considered sensitive data, how they should keep their data secured and trustable and lastly how to verify integrity of their data in cloud computing. This is done by using created data dictionary table where sensitive data are being stored and data control table with hash values. During integrity check the hash value stored in the data dictionary table and with the concatenation of all hash values stored in data control table are compared, if they match it shows that the data has not been tampered with if not something has gone wrong. It did not provide an automatic means of recovering the original copy of corrupted data.

Rady et al (2019) proposed a system that can handle data security problems including, data confidentiality, data availability, data privacy, query integrity verification and query processing over encrypted data. The authors used trusted third party (TTP) server to verify data integrity as it reduces the computation and the communication complexity and cost on the data owner and the user servers. The architecture consists of two phases, comprising the database pre- processing, outsourcing and the user authorization. It also has audit and result phase, which contains query pre-processing querying and result integrity checking. The data maybe compromise in the process of moving data from one person to another.

BosunTijani et al (2021) worked on improving Data Integrity in Public Health: a case study of an outbreak Management system in Nigeria. The article showcases how the design process sparked the concept for intervention to improve the integrity of public health data being collected. The completeness and accuracy of data in the Nigeria health care system is a challenge. Studies have shown that the data quality and extension data integrity has been suboptimal and this poses a barrier to strengthening service delivery. In their work, they ensure that (1) almost all data collection by the test center was now automated, thereby minimizing the proportion of accurate and repeat entry in comparisons of data collected in other part of the same center. (2) Auto validation feature of the system to ensure that all required fields of patient's information were completed and verified. (3) Validations and verification feature to ensure that patients contact information was validated. Limitation of the work was that they did not put any security measure to prevent unauthorized access of patient's data.

The work reviewed so far, did not carry any analysis of the methods used, which would have been a watch dog to any person that want to use it in data processing. In this paper our key interest is on the analysis of hybrid data integrity protection techniques which are limited to validation, digital signature and encryption.

Methodology

In data validation the following rules need to be verified and errors corrected at the input time or at data capture time to ensure that the data meet the rules from the beginning as follows;

1. Does the data supplied have the same type as that specified in the database structure? For example, if the field is alphanumeric, the data supplied for that field must be alphanumeric. If the data field is numeric, the data supplied must be numeric and not alphabetic, and so on.
2. The field width of the data must not exceed the field width defined for it in the data structure. Such mismatches should be flagged and corrected by the appropriate authority in order to avoid automatic truncation in an arbitrary manner that might not satisfy the person that supplied the data.
3. Where possible, value – range checks should be carried out to flag data out of range found in the input source.
4. A standard date field structure must be maintained throughout the date field. One cannot go from dd - mm - yyyy to mm – dd - yyyy or to yyyy – mm – dd or to dd – mm – yy, etc. where dd refers to day of month, mm refer to month of year and yyyy refers to year in question eg 2019. The date field must be further validated to flag invalid dates which conform to the database structure. For example, in the case of dd – mm – yyyy structure, the following are invalid dates: 30/02/2019, 31 – 06 – 2018, 29 – 02 – 2005. This is because 30/02/2019 means 30th February 2019 whereas February can only have a maximum of 29 days in a leap year, otherwise the maximum number of days in February is 28. Also 31 -06- 2018 is wrong because June can only last for 30 days and has no 31st day. 29 - 022005 is wrong because 2005 is not a leap year and therefore February cannot have 29 days in that year.
5. Interfield cross – checks are also important as the date of birth cannot be later than the date of employment. These Validations ensure that only accurate and useful data are being captured and stored in the database for later use.

Pros and cons of data validation

The act of validating data helps organization to check that their databases are correct and valid and this will enable them to make better decisions. in spites of its uniqueness of ensuring that only accurate data is captured right from the beginning it still have some pros and cons associated with it. Here are the pros and cons of data validation.

Advantages of data validation

1. It ensures that stored data at captured level followed the defined rules.
2. In checking the data's accuracy, validating data does a lot of the heavy lifting to ensure data integrity. Validation will not change or improve your data, but it will ensure it serves its intended purpose if it is set up correctly making sure that only accurate data is accepted right from the beginning.
3. It helps Manage Multiple Data Sources. Data validation becomes increasingly important as the number of data sources increases. Suppose one is importing customer data from different channels, there is need to validate all of this data simultaneously against the same tracking strategy. Otherwise, conflicts and errors could appear between the datasets.
4. It saves Time, validating data takes time, but once it is done, there will be no need to change anything until inputs or requirements change.
5. It helps Manage Multiple Data Sources. Data validation becomes increasingly important as the number of data sources increases. Suppose one is importing customer data from different channels, there is need to validate all of this data simultaneously against the same tracking strategy. Otherwise, conflicts and errors could appear between the datasets.
6. It saves Time, validating data takes time, but once it is done, there will be no need to change anything until inputs or requirements change.

Disadvantages of data validation

1. Validation of data is a complex and tough task when several complex data sources are involved.
2. Validation can lead to errors, not all validation software is perfect. Almost certainly, there will be validation errors that need to be fixed**.**
3. One of the biggest problems with validating data is that it needs to be re-validated after certain changes are made. Schema models and mapping documentation must be updated as data types and inputs are provided.

Digital Signature

A digital signature is an encrypted file that travels with the electronic document that needs to be signed and returns with it after the transaction has been completed. Digital signatures are among the most important components of an e-signature program, and they can drive security, legal validity, and records management efficiency when using an electronic signing method. Electronic signatures are making life easier for records management leaders, human resources employees, and managers in a variety of industries. However, digital signature solutions have not become pervasive everywhere, but the foundational technology is in place to drive innovation in even some of the strictest business sectors. Digital signature is a digital code which is authenticated by public key inscribed which is attached to an electronic transmitted document to verify content and sender's identity. The digital signature is the technique for approving legally the authenticity or virtue of message or documents. Digital Signature is equally valid as well as legal as the self attested or handwritten signature. Here are three key reasons why protections of digital signatures are so important:

1. Protecting the signature at the point of signing

The file contains and captures metadata about where the electronic document traveled, which accounts opened it, the IP address of the devices that signed it, the precise timing of the interaction, and other key information. All of this data protects the validity of the signature. With a pen and ink signature, handwriting serves as means of identification. For especially important documents, one can bring in a notary and collect identification for all parties involved. These features protect a physical signature. Similarly, a digital signature that captures the device that an electronic document is signed on, records the user credentials of the person signing the form and captures the pathway that the data traveled between destinations provides multiple tiers of signature verification, protecting the signing process in entirety.

2. Protecting the signature in storage

An electronic record needs to be stored for a significant amount of time, with the duration varying based on specific industry laws. It is often important to verify the stored signature to ensure that it has not been changed over time. A digital signature that is attached to the electronic signature features underlying technology that will show whether the form has been tampered with. It is, essentially, a digital watermark on the e-signature that verifies that the transaction has been completed and seals the file. Anything that breaks that seal is recorded in the digital signature, making it extremely difficult to tamper with the electronic file.

3. Protecting the signature on mobile devices

Smart phones and tablets are becoming pervasive technological tools, and as far as convenience is concerned, they are perfect for e-signing. However, they do not feature the security and data protection features needed to provide the same level of user authenticity as traditional PCs. A digital signature works around this problem by collecting critical metadata in a customizable format. If one need a digital signature to gather more user authentication data to verify a person using a smart phone, it can be configured to perform that task. The result is an operational climate in which one can safely have users sign electronic forms on a mobile device without creating any risk.

Pros and Cons of Digital Signature

The following are the advantages of using digital signatures:

- A digital signature provides better security during transaction. The use of digital signatures and electronic

documents reduces risks of documents being intercepted, read, destroyed, or altered while in transit. An unauthorized person cannot do fraudulence activity with information in transactions when it is digital sign. It is not possible to copy or change the document signed digitally.

- A digitally signed document can easily be tracked and located in a short amount of time.
- Using postal or courier services for paper documents is much more expensive compared to using digital signatures on electronic documents.
- High speed up document delivery, Businesses no longer have to wait for paper documents to be sent by courier. Contracts are easily written, completed, and signed by all concerned parties in a little amount of time no matter how far apart the parties are geographically.
- An electronic document signed with a digital signature is authentic can stand up in court just as well as any other signed paper document, it is 100% legal.
- Signing an electronic document digitally identifies one as the signatory and that cannot be later denied.
- Digital signature is time-Stamp.By time-stamping one will clearly know when the document was signed, date and time are automatically stamped on it.
- No one else can forge digital signature or submit an electronic document falsely claiming it was signed by someone else.

Disadvantages of digital signature

Just like all other electronic products, digital signatures have some disadvantages that go with them, these include:

- There are many different digital signature standards and most of them are incompatible with each other and this complicates the sharing of digitally signed documents. There is need to troubleshoot all the compatibility problems.
- Software is one of the main issues while using a digital signature certificate. To work with digital certificates, senders and recipients have to buy verification software at a cost.
- In some states and countries, laws regarding cyber and technology-based issues are weak or even non-existent. Trading in such jurisdictions becomes very risky for those who use digitally signed electronic documents.
- In order to effectively use digital signatures, both senders and recipients may have to buy digital certificates at a cost from trusted certification authorities.
- Digital signatures, like all technological products, are highly dependent on the technology it is based on. In this era of fast technological advancements, many of these tech products have a short life spam.

Encryption

The word encryption is commonly used to describe techniques that are intended to protect information from unauthorized access. Various methods have been used to transform readable content known as plaintext into a non-readable form known as cipher, from ancient times. The complexity of encryption methods has been successively increasing ever since. Today's cryptographic methods protect data with the aid of computers using very long key lengths, complex mathematical tasks, and sophisticated key exchange procedures. As a vital element of information security, encryption helps to secure electronic communication between people and devices from spying and manipulation. Encryption is a method commonly used in IT to protect data from unauthorized access. To do this, algorithms turn the data from a readable into a non-readable format. The information is now only accessible to those who possess the keys needed to decrypt the data.There are two forms of encryption, symmetrical and asymmetric. But there are several types of encryption as well as different algorithms to aid the encryption process. The benefits of encryption include increased security and privacy, but a disadvantage is that this system requires detailed planning and maintenance.

Although it seems like common sense to use data encryption in business and other entities for security, many organizations are opposed to encrypting data because of some of the obstacles involved with doing so.

Advantages Data Encryption

- Data encryption allows the data to remain separate from the device security where it is stored. Security is included with the encryption which permits administrators to store and transmit data via unsecured means.
- Data encryption circumvents the potential complications that accompany data breaches which provide ensured protection of intellectual property and other similar types of data.
- Because encryption is on the data itself, the data is secure regardless of how it is transmitted. An exception to the rule can be transmission tools such as email because sometimes a typical email account does not provide the necessary security.
- A lot of organizations are required to meet specific confidentiality requirements and other associated regulations. Encrypting data means that it can only be read by the recipient who has the key to opening the data.
- Encrypted data maintains integrity. Integrity describes data that is kept complete, accurate, consistent and safe throughout its entire lifecycle. And encryption technique guarantees data integrity

Disadvantages associated with Data Encryption

- Without a doubt, data encryption is a monumental task for an IT specialist. The longer data encryption keys the more difficult IT administrative tasks for maintaining all of the keys can be. If you lose the key to the encryption, you have lost the data associated with it.
- Expense data encryption can prove to be quite costly because the systems that maintain data encryption must have capacity and upgrades to perform such tasks. Without capable systems, the reduction of systems operations can be significantly compromised.
- It is have unrealistic requirements. If an organization does not understand some of the restraints imposed by data encryption technology, it is easy to set unrealistic standards and requirement which could jeopardize data encryption security.
- Encryption is not a compatibility technique. Data encryption technology can be tricky when layering it with existing programs and applications. This can negatively impact routine operations within the system.

Results

Table 1, summarizes the encryption result with respect to the integrity, cost etc.

Table 1: The techniques analyzed are summarized in the table below:

Summary of the Analysis			
	Validation	**Digital Signature**	**Encryption**
Integrity	Validated data ensures that accurate data are accepted from the beginning, thereby maintaining the integrity of data.	It make sure that the validated data stored with digital signature remain unchanged throughout its life span	Encrypted data maintains integrity of data
Cost	Validation software is affordable	It is less expensive	Data encryption can prove to be quite costly
complexity	Validation is tough with complex data sources		In encryption there is unrealistic requirements , and this could jeopardize data security
Security		It provides better security during transaction or even when the data is at rest , thereby making it impossible for any fraudulence activity	Encryption secures data regardless of how it is transmitted
Acceptance	It is highly accepted by any organization that uses data for decision making	It is highly accepted	Encryption Technology is widely accepted
Data Management	It helps Manage Multiple Data Sources		Encryption allows data to remain separate from device security
Time	Validation ensures that there is no need for change anything until inputs requirements change thereby saving time	It can easily be traced and locate in short amount of time	It can easily be traced and located in short amount of time
compatibility	No compatibility problem invalidation	There are many different digital signature standards and most of them are incompatible with each other and this affect sharing of digital signature	There are compatibility problem associated with encryption. It can be tricky when laying with existing program and application

Conclusion

Analysis of Hybrid Data Integrity Protection Techniques was done in this paper, where the pros and cons of each technique was stated and summarize to give a better picture of each method. This should save as a guide in using any of it for data integrity protection. Ensuring that data maintain its integrity through out there life span is very important, and using hybrid techniques provides levels of protection to the data.

Reference

1. Bosun T, Toni J, Busayo O, Zahra K. Improving Data Integrity in Public Health: A case Study of an outbreak Management System in Nigeria. Glob Health SciPract,2021:29-9(2):S226-S233. published online
2. Priyanka D, Tushar C. Hybrid Encryption for Database Security International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, 2017, 04(11). www.irjet.net p-ISSN: 2395-0072
3. Mai Rady, Tamer Abdelkader, Rasha Smail. Integrity and Confidentiality in Cloud Outsourced Data. Ain Shams Engineering Journal,2019:10(2):279–285.
4. MrinalK, Sanjay K. Ensuring data storage security in cloud computing based on hybrid encryption schemes. Parallel, Distributed and Grid Computing (PDGC), 2016 Fourth International Conference on, 2016, 22-24.
5. Chinnasamy P, Deepalakshmi P. Design of secure storage for health-care cloud using hybrid cryptography. In: 2nd international conference on inventive communication and computational technologies (ICICCT 2018). IEEE Xplore Compliant-Part number: CFP18BAC-ART; ISBN 978-1-5386-1974-2, 2018.
6. Lawrence Hart. Pros and Cons of electronic Signature, 2022.