iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** VII **Month of publication:** July 2024

DOI: https://doi.org/10.22214/ijraset.2024.63767

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VII July 2024- Available at www.ijraset.com

Data Leakage Detection Using Cloud Computing

Sakshi Rahangdale[1], Pallavi Bansod[2], Nikita Bante[3], Radha Yete[4], Priyanka Jagtap[5]
Professors, Department of Basic Science & Humanities, RTMNU University Nagpur, India

Abstract: ***The exchange of data between users is crucial in the current environment. Distributors, who are typically the data owners, send data primarily to users who are interested in information from reliable third parties. The distributor must transmit information in a secure manner and with confidentiality. Data leaking is the term used to describe the situation where some parties create duplicate copies of the same information during data sharing, resulting in significant losses. One must take steps to identify the data leak early on in order to stop it from happening. This study discusses data leakage and how different watermarking approaches can prevent it.***

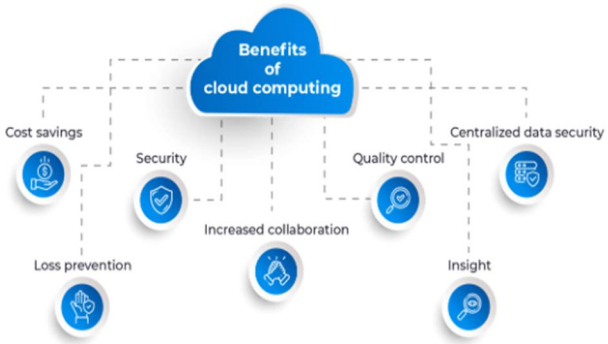
Keywords: Data leakage, resilient watermarking, discrete wavelet and discrete cosine transforms, and watermarking.

I. INTRODUCTION

Virtually every IT company is making an effort to get into the cloud computing market, which is one of the fastest growing and most promising technologies in the market today. The newest method of obtaining delivery models for the IT business is emerging called cloud computing. IT services can be delivered in the form of SAAS, PAAS, or IAAS using this method.

Providing internet services like networks, storage, servers, and software—which are essentially on demand—can also be done via the cloud. By enabling pay by perusing, or pay according to user consumption, it reduces hardware costs and saves time. Put differently, cloud computing is a conglomeration of cutting-edge platforms and technology that offer online hosting and storage services. Cloud computing offers high-quality on-demand infrastructures at a lower cost. The following are the primary benefits of cloud computing:

1) It reduces hardware and maintenance expenses.
2) It is adaptable
3) Is accessible from anyplace in the globe
4) There is no need to worry about software upgrades because it is fully automated.
5) Data storage off-site.
6) Help during a disaster
7) Always up.



In order to store data in the data centers owned by the service providers, cloud technology is entirely dependent on the internet. Data security is one of the key problems with cloud computing technologies. Less control over data can result in major security risks and problems, including data leaks, data vulnerability, and outside or insider attacks on data. In order to prevent data leaks, every IT firm must focus on security issues connected to shielding their data from various third parties. Sometimes insiders, primarily current workers of the organization, leak company secrets. In these cases, the security must be unknown to the employees, preventing them from knowing how to break it. Data breaches are unpredictable and can happen at any time. The caliber of sensitive information disclosed by the leaker alone determines how much harm is caused by the data leak.

Individual if the information that has been disclosed is crucial to the organization. It might render the organization defenseless. Leakage may cause the firm to decline and ultimately lead to the company's demise.
The Watermark approach is one of the many methods for preventing data leaking that have been developed to address this problem.

II. METHOD OF WATERMARKING

A watermark prevents the data owner's copyright from being violated. It's a method where every distributor copy has a special code included in it. Basically, it is encryption applied to a specific piece of data that needs to be shared. Images, videos, or any other type of significant file can contain information. The watermark facilitates the company's ownership claim over certain data.
This approach adds a small pattern to the data, which are mostly the tuples and the subset of data. Only the data owner can access the tuple and subset's properties because they are algorithmically coded and controlled by a private key. The watermark is represented by this design. Only when someone obtains the key can the data be accessed.
It is not necessary to have access to the original data in order to identify the watermark. If the data includes some of the markings, Even with a small sample size of the data, the watermark can be distinguished. The watermarking is carried out using software that uses watermarking techniques to embed watermarks. A few data inaccuracies are introduced by the software. These mistakes are called markings, and a watermark is made up of all these marks combined.

III. VARIOUS WATERMARKING METHODS

1) Applying watermarking using Discrete Cosine Transforms (DCT)

A discrete cosine transform, or DCT, is a method for dissecting a signal into its fundamental frequency components. The image is initially transformed into 8x8 pixel blocks using this method. Following DCT conversion, a Gaussian network classifier is used to identify the mid-frequency range. The mid-frequency DCT coefficients are now employed in the embedding process. A linear DCT constraint modifies the DCT coefficients. The image's visibility won't be impacted, and compression won't get rid of the watermark.

2) Discrete Wavelet Transform (DWT) watermarking technique

This is a modern method that's widely used for watermarking, compressing pictures, etc. This technique uses wavelet filters to modify the image. Wavelets are short waves that occur at several different frequencies. The wavelet transform splits the image into three spatial directions: diagonal, vertical, and horizontal. The core concept of DWT is multi-differentiate decomposition, which divides a picture into sub-images with independent frequencies and unique spatial domains.

3) The least significant bit, or LSB

This method incorporates the watermark into the LSB of the pixels. Despite being easy to use, this approach is not very secure against assaults, as it is easy to delete the watermark. To watermark a picture, choose a subset of pixels, and then replace each selected pixel's lower-left corner (LSB) with a watermark bit.

4) Wetmarking via Integration and Removal

Using this technique, a watermark is created by encoding a small portion of the main cover image's fractional pixel intensity value. The image's accuracy is preserved by the watermark in the unimportant. watermark is undetectable when using this technique. Businesses and corporations who provide digital information security items would benefit from this technology, which makes it easy to implant and retrieve a substantial amount of data. It is an added advantage of this strategy. This technique makes use of many embedding and extraction techniques.

5) Watermarking using Wavelets

This technique embeds the multi-resolution data fusion, using discrete wavelet form transformations for the watermark and picture. The watermark is a part of every wavelet level. The average of the guesses from each wavelet decomposition resolution level is used to determine the watermark. This algorithm can be used for filtering, additive noise reduction, and JPEG compression.

6) Watermarking using Secure Spread Spectrum.

A watermark needs to be added to important signal components in order to conserve multimedia data, including music, video, and images, and to make the signal resistant to malicious attacks and common signal distortions. However, the data signal may deteriorate as a result of these component alterations.

Therefore, using techniques similar to spread spectrum communications, a watermark must be added to the spectral components of the data in order to hide a narrow band of the signal within a wide band signal. Even if multiple outsiders collaborate with distinct versions of the watermarked material, it is exceedingly impossible for an external to erase this watermark.

7) Sturdy Watermarking Method

This watermarking technique is immune to typical operations such as compression, noise addition, filtering, and so on, as well as geometric attacks such as rotation, scaling, translation, and so forth. It's commonly used to protect ownership. In the event that ownership rights are disputed, the data can be erased and used to identify the legitimate owner by explicitly encoding ownership information into the data in this way. The watermark ought to hold up during extraction and not cause any damage to the data.

8) Digital audio and image watermarking using Mat Lab

Using the RSA algorithm, a watermark is encrypted and then inserted using the LSB approach into the audio file. The outdated LSB technique is not very effective in fending off attacks. This makes it extremely difficult to remove the watermark since the original watermark is encrypted before being included into the audio file. Very high robustness results from this. The embedded watermark is initially recovered and then decrypted during data retrieval. Similar to this, the DWT approach is applied to image watermarking. As a pseudo-noise sequence, the watermark is embedded. This technique removes the embedded watermark from the watermarked image or audio while maintaining the highest level of security for the image and audio. The original watermark must be known.

9) Watermarking without being seen

This technique offers a strong, invisible watermarking strategy that may be used to extract and embed a digital watermark into a picture. This technique uses a sub image as a watermark, which is one of its main features. The quality of the image will suffer if the watermark is changed because it is inserted in the most crucial region of the original. Two stages go into creating the watermark. Synthesizing an image from its sub-image is the initial step in the process. The next step involves applying a visible watermarking technique to the synthetic image in order to include a logo, or watermark. This creates a compound watermark. Next, the primary block of the host image has this compound watermark subtly incorporated into it.

10) Data security is necessary

The information that makes up the data is mostly private and sensitive information. The data may have come from a basic person, a large IT company, or a company. The owner is not protected if these data are disclosed. Uncontrolled data has the potential to make a big organization insecure. Cybercriminals, or the person who stole the information, lose money, an organization's reputation, its brand value, and the confidence of its customers when they sell it for profit. Therefore, in the modern world when all information is digitized, data security is essential.

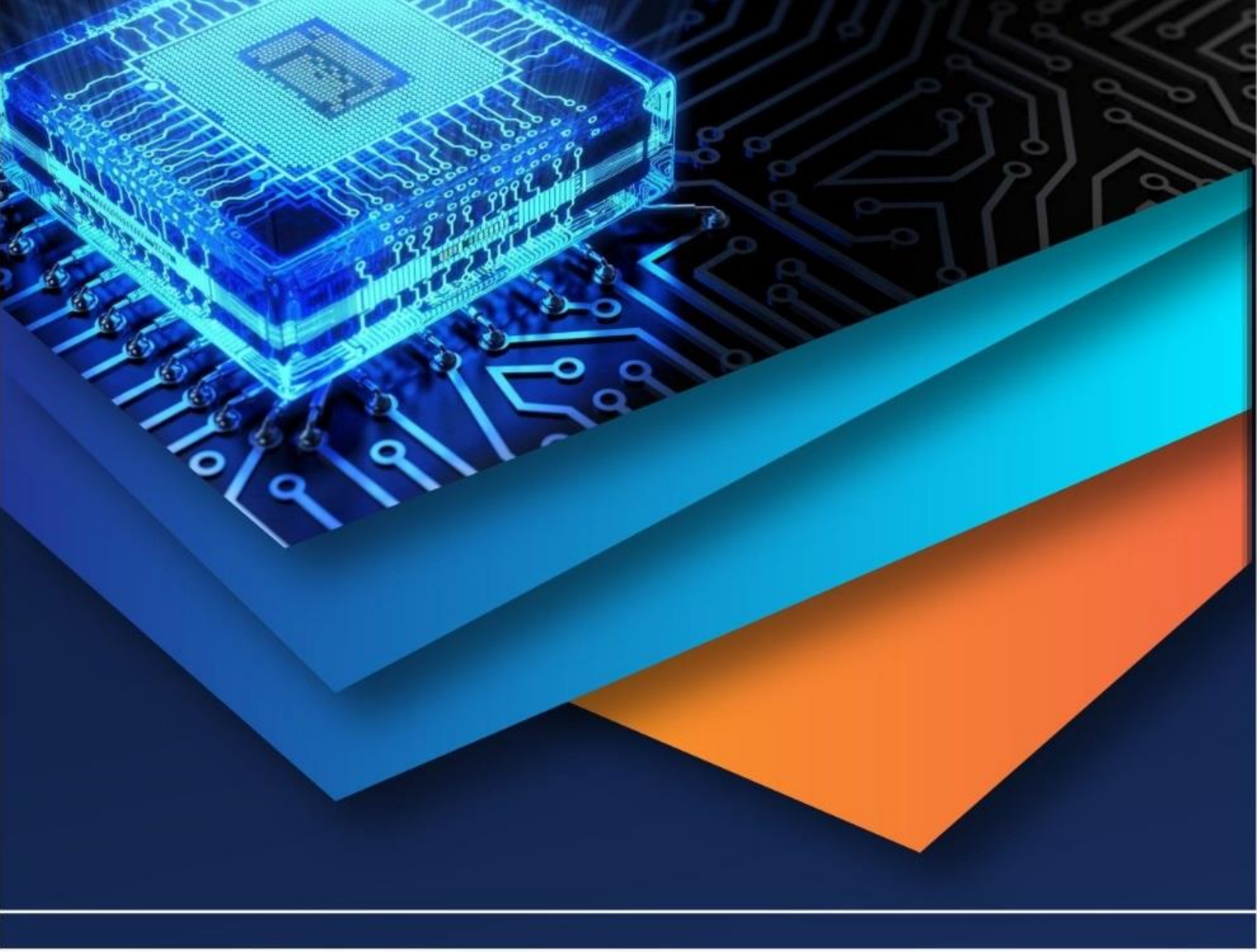
IV. CONCLUSION

One kind of silent yet deadly hazard is data leakage. Without awareness, sensitive information can be disclosed. It could be the work of an outsider or an insider. Therefore, before being distributed, sensitive data needs to be watermarked so that its origin can be determined with complete certainty. When data is watermarked, it prevents unauthorized access and makes it simpler to identify the guilty by employing fictitious objects, such as watermarks positioned at various informational locations. This study covered a variety of watermarking techniques and their significance for data security, including DCT, DWT, wavelet, invisible, etc.

REFERENCES

[1] P. Papadimitriou and H. Garcia-Molina, “Data leakage detection,” IEEE Transactions on Knowledge and Data Engineering, pages 51- 63, volume 23, 2011.
[2] IEEE Transactions On Knowledge And Data Engineering, Vol. 22, No. 3, March 2011 Data Leakage Detection Panagiotis Papadimitriou, Member, IEEE, Hector Garcia-Molina, Member, IEEE P.P
[3] An ISACA White Paper Data Leak Prevention P.P
[4] Ensaf Hussein, Mohamed A. Belal, “Digital Watermarking Techniques, Applications and Attacks Applied to Digital Media: A Survey”, IJERT, ISSN: 2278-0118, Vol. 1 Issue 7, September2012.
[5] Sandip A. Kale, Prof. S.V.Kulkarni, “Data Leakage Detection, International Journal of Advanced Research in Computer and Communication Engineering”, ISSN: 2278- 1021, Vol. 1, Issue 9, November 2012.
[6] Upasana Yadav, J.P.Sharma, Dinesh Sharma, Purnima K Sharma, “Different Watermarking Techniques & its Applications: A Review”, IJSER, ISSN 2229-5518, Volume 5, Issue 4, April-2014.

[7] Cox, I.J.; Miller, M.L.; Bloom, J.A., "Digital Watermarking", Morgan Kaufmann, 2001.
[8] Prabhishek Singh, R S Chadha, "A Survey of Digital Watermarking Techniques, Applications and Attacks", IJEIT, ISSN: 2277-3754, Vol. 2 Issue 9, March-2013.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY