Word Template for Term (Survey) paper

Your Name
Your affiliation

Your E-mail

ABSTRACT

Combining several media formats, including text, audio, images, and videos, into one digital file is known as multimedia. It's a kind of communication where messages or information are spread through a variety of mediums.
Many industries, including education, entertainment, advertising, and communication, can benefit from the usage of multimedia. Multimedia has become an indispensable aspect of our everyday existence, and as technology has advanced, it has grown increasingly dynamic and captivating.

Keywords

Multimedia, Text, audio, Captivation, Steganography, Cryptography, Big Data, Integrity, Confidentiality, Encryption, Decryption

1. First Page Copyright Notice

2. Introduction

The term "multimedia" describes the use of several media to deliver a message or convey information, including text, audio, images, animations, and video. In order to give the viewer a more varied and interesting experience, it entails integrating different media types.

Multimedia Systems are flexible tools for a wide range of applications because they are designed to store, retrieve, and analyze such a vast range of information efficiently.

Enables users to engage with information in many formats is a key objective of multimedia systems. They make it possible for people to produce, modify, distribute, and exhibit material in a multitude of ways, which promotes effective expression and communication.

Consider a multimedia presentation that tells a gripping story or presents a difficult message by skillfully fusing text, graphics, videos, and audio. Multimedia systems, which offer the underlying infrastructure to manage these many elements with ease, enable such presentations.

In current era of ubiquitous information devices (Smartphones, tablets), the consumption of digital media and its influence on decision-making processes (e.g., elections) has achieved a majority-owned relevance over conventional media (e.g., printed newspapers). We also have to understand that there will inevitably be a loss of data authenticity in the consumed information along with that cultural shift.

Misinformation and disinformation are more common on these platforms than in traditional media since there is inadequate fact-checking and third-party screening. Even when people are presented with real facts, the dissemination of false content can have a lasting effect on their attitudes.

3. Development of Big Data Security for Multimedia

A new paradigm called "big data" is being used to datasets that are greater in size than what can be captured, managed, and processed with software alone necessary time period. These datasets often come from a variety of sources, both of which are massive. Big data is defined as the gathering of the three values—Volume, Variety, and Velocity—and Specific technology and analytical methods are required for using inference to turn the data into knowledge mechanism it into a valuable product.

Multimedia data is growing at an exponential rate, presenting both new security issues and previously unheard-of benefits. Big data that is multimedia, comprising text, pictures, audio, and video, is becoming more and more significant in a variety of industries, including social media, healthcare, education, and entertainment. However, there are substantial security problems due to the sheer amount, diversity, and velocity of multimedia big data.

The exponential growth of multimedia data presents unprecedented advantages as well as new security risks. Multimedia big data, which includes text, images, audio, and video, is going to be important in a lot of different areas, such social media, healthcare, education, and entertainment. However, the sheer volume, diversity, and velocity of big data in multimedia poses significant security challenges.

But the sheer amount, diversity, and speed of multimedia big data present serious security issues, such as:

Confidentiality: It is essential to guard sensitive multimedia data against illegal access in order to stop illegal viewing, alteration, or destruction.

Integrity: To preserve confidence and avoid manipulation or tampering, multimedia data must be guaranteed to be authentic and in its original condition.

Availability: Ensuring that authorized users can access multimedia material when needed is essential for both user happiness and company survival.

Privacy: In order to abide by laws and uphold public confidence, it is crucial to protect the privacy of people and organizations that handle multimedia data.

Every day, social media sites, e-commerce sites, sensor networks, the health sector, and banking systems generate enormous amounts of data. Data was once measured in Kilobytes, but in the current environment, data is measured in Petabytes or ExaBytes, which is causing the data life cycle to get more complex every day. Big data multimedia

offers a large amount of space for this data's collection, processing, storage, and distribution.

One can extract information (at second level) from vast amounts of unstructured and semi-structured data using a variety of big data tools and techniques (at first level). This information can then be utilized to depict knowledge (at third level) in accordance with requirements, and decisions (at fourth level) can be made based on the knowledge gained in addition to human input.

Researchers and practitioners are creating a variety of security solutions for multimedia big data to tackle these issues. These remedies may be divided roughly into three categories:

Cryptographic Techniques: To safeguard the confidentiality and integrity of multimedia data, cryptographic techniques like encryption and digital signatures are used. Multimedia data is rendered unintelligible by encryption techniques, yet digital signatures offer an impenetrable means of confirming the data's legitimacy.

Access Control methods: Access control methods govern who has access to and may operate on multimedia data. Examples of these mechanisms include role-based access control (RBAC) and attribute-based access control (ABAC). These safeguards prevent unwanted access and guarantee that sensitive data may only be accessed by those who are authorized and have the necessary rights.

Privacy-Preserving Techniques: These methods, which include differential privacy, anonymization, and pseudonymization, are used to preserve people's privacy while allowing for data analysis and the extraction of valuable insights. These methods strike a compromise between the requirements for data usage and protection.

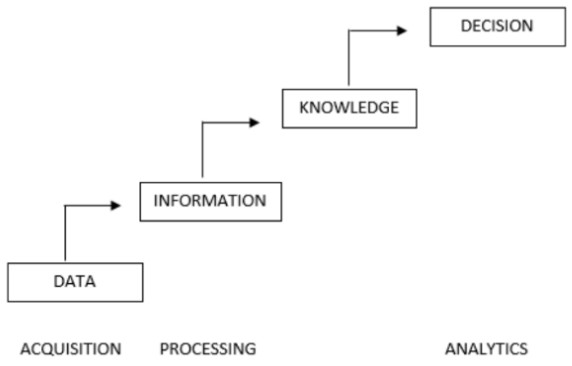

Figure 1: A Typical Life Cycle of Multimedia Computing

Problems and Difficulties Many issues are emerging as a result of the massive volume of data being generated. Issues with multimedia big data security may be regarded from several angles since there are numerous obstacles, ranging from data filtering to data decisions:

a) **Various Data Formats:** Three forms of data are found in multimedia big data [24]:

i) Organized

ii) Unorganized

iii) Relatively unstructured

b) **Scalability:** Multimedia big data has a vast user base that produces massive amounts of data through the use of a multitude of devices across many platforms on a regular basis, necessitating extensive processing in order to optimize storage and transmission resources. A wide range of platforms and devices on a daily basis, requiring a great deal of processing power to maximize storage and transmission capacity.

c) **Multi-level Protection**: Since sensitive information is included in multimedia at every stage, protection should go beyond file level protection and include access control methods that may offer security rules for batch and stream processing. All of the personal data is contained in multimedia, together with the rights of access to it. Thus, protection needs to be provided at all levels.

d) **Data Acquisition:** Humans are creating multimedia data from several sources, including machineries in substantial portions. Social networking, health, and wearable sensor platforms are all contributing to the increasing amount of data flowing into the streams. It is possible to forecast a person's health state using data obtained by their daily use of their mobile phone. Numerous pieces of information pertaining to specific individuals are gathered during the procedure. It is therefore a difficult undertaking to keep such data in a secure and organized manner. There ought to be methods for optimizing the data so that it can be presented clearly and that appropriate conclusions can be drawn.

e) **Accessibility:** When a user attempts to access his data using various devices and from any place at any time, it should be available to him. Thus, data access management would be in place to make it simple to retrieve relevant data from any place that has the appropriate authorization and authentication. Moreover, role management is crucial to Data access management may make data easily and conveniently accessible, ensuring that users won't experience any difficulties when transferring. Therefore, a technique for locating users should exist, together with the possibility that their location won't impact data access procedures.

f) **Data Compression/Reduction:** There is a significant quantity of semi-structured or unstructured data combined with characteristics like redundancy. When it comes to noise and inconsistency, they should be minimized to the extent necessary for efficient storage and communication. Effectively with a small

amount of computer resources. There are several methods for compress data, such as by RED encoding, AST (Abstract Index Trees), sensing, and spatiotemporal data. These methods allow for the compression of an image or video together with its attributes, which may then be applied to the picture or video's computing process. The majority of data generated nowadays is stored on cloud servers, which provide consumers flexible storage at extremely cheap cost. Users create a great deal of unstructured data in the cloud, which may be reduced or compressed using various compression techniques before being stored. One of the biggest challenges facing the multimedia big data community is applying such approaches without sacrificing data quality.

g) **Data Representation:** Since data is available in a variety of forms, including text, images, audio, and video, and is present in huge chunks that are either semi-structured or unstructured. It must be analyzed in a way that allows for appropriate representation and the extraction of some information from the data. It is a major effort to separate the data in a stream processing and to describe this data since there should be some system that can represent data before and after analysis.

4. Multimedia Based Steganography

4.1 What is Multimedia based Steganography

A method for embedding hidden messages into multimedia data, including music, video, and graphics, is called multimedia-based steganography. Multimedia-based steganography aims to conceal the secret message's presence so that unauthorized parties are unable to decipher it. The information in Steganography can be hidden or concealed using a variety of techniques. By incorporating encrypted material through audio and video files, it provides increased security. As seen in Fig. 1, there are several ways to conceal hidden material in multimedia-based steganography files, including text, picture, audio, and video files. The volume of the information stays the same once it has been hidden in the multimedia-based steganography file, which is the fundamental benefit of embedding safe data in these types of files.

***3.1.1 Steganography of Text*:** The information is encoded within the text itself. The nth character in a sentence of the document's content hides the system's crucial information. In a written document, there are several ways to hide data.

They are listed in the following order:

i) Method Based on Format

ii) Methods of Random and Statistical

iii) Linguistics Method

The goal of this work is to maximize text Steganography's utilization capacity, or text capability, by employing ASCII values. Using pixel values and ASCII values, a novel method has been presented to encrypt the text or content inside the photographs.

The core of this approach is the merging and division technique, which yields the outcome. With this method, pixel values are given more weight. This model, in contrast to previous methods, may encode or embed all 255 ASCII letters; nevertheless, it is only compatible with lossless text formats. Ultimately, it can be concluded that this method effectively uses steganography while meeting its requirements. With a payload size of 10 Kb and a PSNR value of 44.0, the outcome is rather respectable.

The author of [3] proposes a contemporary solution to format-based content steganography by combining a copy protection method with two content or steganography techniques: line-shifting and character or word-shifting methods. By using the aforementioned techniques, mask the information in binary or digital format rather than character format. minimal than one bit is concealed in a line of the cover text, indicating that this approach has a significant hidden ability and produces minimal distortion in the cover text. This method is hybrid as it uses a particular character to conduct text steganography and combines two techniques (word shifting and line shifting).

According to the author of [2], steganography is a technique for hiding messages in such a manner that its presence cannot be taken into account as an improvement in the efficiency of steganography algorithms. The information hiding limit in this study is improved by the provided algorithms, which include the Zero Distortion Technique, random sequence curve, steganalysis, and the matrix of allocation on Text.

Clients of the suggested technique can conceal additional data points without changing the dimensions of the cover image. It suggests that the reflected changes are essentially negligible. The steganographic algorithm's efficiency improvement for greater performance is the paper's restriction. Since word processing software destroys data when spaces are removed, optimization for resilience needs to be enhanced.

3.1.2. Steganography of Images: Steganography of Images By designating the cover medium or item as a picture, image steganography obscures specifics. Consideration is given to pixel quality in picture steganography. Since digital picture representation contains many bits, photographs are frequently utilized as a cover source in digital steganography to encrypt information. With the help of the steganographic Generative Adversarial System (SGAN), several advancements in computer vision and natural language translation have been made possible. Picture Steganography is the process of concealing material within a digital image in order to accomplish clandestine communication. This model concealed and rejected steganalysis. employing the pseudo-random encoding method, which embeds the concealed data in the master document's key and sends it to a designated recipient, together with the LSB algorithm or measurement. The PSNR values of these processes are over 60 dB, according to the results, and the pseudo-random approach outperforms the LSB process (a number generator that generates random numbers quickly).

3.1.3. The process of audio steganography: In this method, binary or digital audio contains hidden material. The audio file's binary categorization provides solutions for audio steganography, which is achieved by making little adjustments to the data.

This involves data being embedded into audio files. With this method, the sound files in mp3, au, and wav are hidden using certain audio steganography techniques. There are several different audio steganography methods. These methods consist of the following:

i) Disperse Spectrum

ii) Phase Coding

iii) Low Bit Encoding

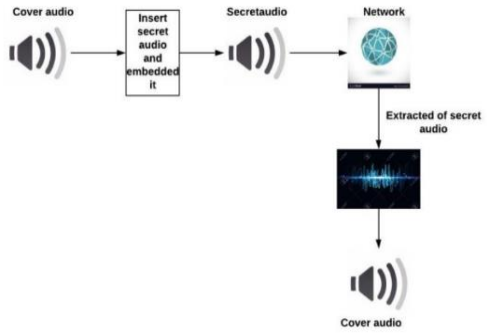

Fig. 2: Audio Steganography (a).Cover audio: This is the normal audio signal that is used to mask the secret audio signal. (b).Secret audio: This is the audio signal that is hidden in the cover audio signal. (c).Steganographic embedding: This is the process of embedding the secret audio signal in the cover audio signal.

We can create an alternative steganography technology to minimize the impact on sound. This is used to bury or conceal information within it. This paper's main goal is to present a practical method for masking or disguising the content and making it a more secure location. The audio file becomes corrupted when the header component is missing.

The LSB algorithm is primarily used to insert or integrate the secret content into a distributed or spread media. Here, the creator uses genetic programming to provide more security to the information so that the secret information cannot be recognizable. In the proposed approach, the secret message is embedded in the sound document as noise and by using general audio decline the dissimilarity between first spread record and Stego document. Audio records are used as spread or cover media, and text content is used as secrete data.

Unpredictable and higher LSB steganography is used to hide the image content in audio samples. Because the audio file on the cover is altered to create a Stego-file by modifying several bits, this method is resistant to Stego analysis attacks. Strong PSNR functionality is present in the related Stego wav file. It is undetectable to add the final audio file because it is perceptually identical to the original. The application of these steganography techniques is challenging to do but also extremely safe and dependable. The author of this work discusses the current method of enclosing the message within frames in order to minimize auditory damage. Without any mistakes, it can recover embedded or buried messages. It has the capacity to repair the concealed or embedded communications flawlessly. The suggested method is effective against some benign assaults. Strategies for hiding must adhere to certain requirements. To effectively cover a lot of information in audio without seriously harming the eyes.

3.1.4 Video Steganography:

This is a method of concealing content in binary or digital video format. Video (consists of a set of images) is usually used as the cover file for embedding the contents. Typically by using (DCT) i.e. discrete cosine transform algorithm modify the values (9.667 to 10) that are used to mask the material in each entity of the images in the video that the human eye cannot identify or locate. In video Steganography the formats used are MP4 and MPE are used. In all of these approaches the main motive is to conceal a secret message in another cover item. This may not be able to guess the presence of information inside cover media. It is necessary to apply the proper decoding method to retrieve the information.

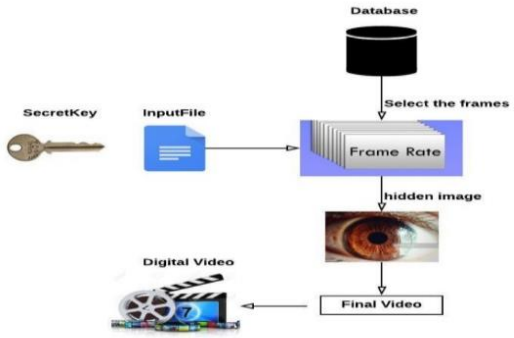

Fig. 3: Video Steganography (a). Select the frames: This involves choosing the individual images that will make up the video. The frame rate is the number of frames per second that will be displayed, and it affects the smoothness of the video. (b). Encode the frames: This involves compressing the frames into a format that can be stored and transmitted efficiently. (c). Combine the frames: This involves assembling the encoded frames into a video file.

For the above diagram, This is a technique that encrypts a document or other sort of content in video format using the video as a cover file. In order to conceal the information that is visible to human eyes in each picture in the movie, the Discrete Cosine Transformation (DCT) typically modifies the values. In the study, the embedding payload will be modified utilizing alternative techniques to get higher video quality.

In [7], Instead of classifying the smooth sections of the frame using Grey Correlation, the author proposed utilizing ECC (Error Correcting Code) to encrypt the Stego message and then storing the details in the picture being conveyed using the recursive steganography approach.

To address issues and parallelize the process for a more effective video steganography algorithm, the author of [8] suggested employing LSB as an efficient video steganography technique. A symmetric key was used to achieve secure message encryption, and the feedback register (FSR) was used to randomly select the edges or frames for encryption. The proposed solution involves using four strings in a cycle of parallelization, encryption, and encoding/extraction procedures that are carried out in parallel, hence increasing the amount of the details. The algorithm may be improved much more to process data in real time.

The author of [9] suggested combining steganography and cryptography. The system's main goal was to conceal data in the video by using ASCII code encryption and the pre-existing technology. It offers two different levels of defense.

Cryptography provides confidentiality, whereas Steganography is intended to keep information secret. The author of [10] proposed employing a graphical password, LSB method to convert it into a binary form, and DCT and DWT to obtain watermarked images. Using all three of these combinations of techniques raises the security level. The encryption on both approaches is more challenging.

Future work aims to improve the outcome by utilizing an algorithm different from the one employed in the thesis. Table 1 discusses the comparison of steganography techniques based on multimedia.

Techniques	Imperceptibly	Robustness	Capacity	Complexity
Transfer domain	High	High	Medium	High
Spread spectrum	Low	Medium	Low	Medium
Parity coding	High	High	Low	Medium
Echo hiding	Medium	High	Medium	Medium

5. Big Data Encryption Techniques in Multimedia

In this world where data security is everyone's top concern, multimedia big data encryption techniques are necessary at every stage of the data life cycle, including acquisition, storage, processing, dissemination, and presentation. These days, the majority of organizations and users are gravitating toward the cloud to store and process their data because it offers resources on demand in a convenient way and also provides storage at a very low cost. Cloud providers protect the security of stored data by using encryption techniques on them.

Whether the data is in use, in motion, or at rest, multimedia big data encryption techniques are necessary. Encrypting data used for communication purposes is always necessary to prevent man-in-the-middle attacks. Data integrity and consistency should be ensured by the method encryption is applied to data. A vital and significant part of every encryption method is key management.

The decryption key should only be intended for the recipient so that he may gain the required access. Many cryptography approaches are used for key management. Two different cryptographic techniques—symmetric and asymmetric cryptography—are employed for key management.

Unpredictable and higher LSB steganography is used to hide the image content in audio samples. Because more than one bit is manipulated to create a Stego-file from the audio file in the cover, this method is resistant to Stego analysis attacks. The accompanying Stego wav file has a high PSNR performance. It is undetectable to upload the final audio file because it sounds exactly like the original. While extremely safe and dependable, putting these steganography tactics into practice is challenging. In order to lessen auditory damage, the author of this work discusses the current method of hiding the message inside frames . Error-free restoration of embedded or concealed messages is possible.

Different encryption methods are used for various multimedia files. There are various picture file formats, including JPEG, PNG, and others, that require encryption. DES additionally. They are usually encrypted using AES techniques. Additionally, a thumb-nail preservation approach is presented to encrypt the picture files.

In order to provide the user with an accurate resolution preview of the image while maintaining the image's encrypted form, this technology generates correct, reduced resolution thumbnails from the encrypted images. Every encryption method specifies the performance parameters that it uses to encrypt data in any format or multimedia type.

For instance, the user can assess and compare different methods with regard to tenability, visual degradation, compression friendliness and openness, format acceptance, encryption proportion, speed, and cryptographic security for image files.

The performance metrics for various formats could change. Every byte in audio files is encrypted using 3DES or AES, and without a decryption key, it can take years to unlock the file. Various encryption techniques are employed for different types of video encryption. These include permutation-based encryption, fully layered encryption, selective encryption, and perceptual encryption. The primary uses of encryption are various forms of video encryption, including fully

6. Comparison of Methods And Algorithms

It is evident that there is no set algorithm for encrypting multimedia data. DES and RSA are often employed methods for data encryption, either for user authentication, before data is stored, or to create a secure channel for data transmission [5]. To encrypt a substantial volume of data, we employ a hybrid methodology. The word hybrid refers to a combination of two or more methods that use symmetric encryption to protect data techniques (DES and AES), and an asymmetric scheme (RSA) is used to transfer the key.

1) **Content-based image retrieval (CBIR):** The method of obtaining photos from a big database according to their visual content is called content-based image retrieval, or CBIR. CBIR uses characteristics like color, texture, shape, and spatial layout that are taken from the photos themselves rather than keywords or metadata. This enables users to look for images, even ones without accompanying text, that are comparable to a query image.

2) **Bag-of-Words (BoW):** One method for describing text documents as vectors of word counts is the bag-of-words (BoW) model. It is a frequently used technique in natural language processing (NLP) and information retrieval (IR) for efficiently extracting features from textual input. The multimedia BoW paradigm has a number of benefits.

i) **Simplicity:** This strategy is clear-cut and simple to use.

ii)**Scalability:** It can be effectively used with huge multimedia data collections.

iii**)Effectiveness**: It can work well for a variety of multimedia jobs, especially when paired with other strategies.

3) **Video Encryption Algorithms:** As multimedia technology has advanced and the internet has grown quickly, video dissemination has become very easy. As a result, video file security has become crucial. Video files require a very long time to be encrypted using the traditional encryption process; to cut down on this time, they

favor the somewhat insecure scrambling techniques. When it comes to big real-time videos, encryption. There are two categories for algorithms: light weight encryption and selective encryption.

4) **DRM on Identity Based Encryption**: Using cryptographic technology, this is one of the most effective and dependable tools for safeguarding digital content. The IBE is entirely dependent on two factors: first, it makes certificate administration and management simpler by utilizing unique. Along with that, it specifies security by employing an elliptic curve cryptography as a second parameter. Secret data and well-known information are referred to as public key. By utilizing these factors in a unique way, IBE reduces the overall computational cost in comparison to traditional cryptosystems and facilitates key management.

7. Conclusion

This study offers an overview of the different steganography methods, including their forms and classifications, that have been suggested in the literature review over the last few years. Additionally, this study surveys a number of multimedia Big Data security topics. It focuses on the difficulties and problems with multimedia data security as well as the numerous and rapidly expanding multimedia approaches to data security. We have covered a variety of encryption techniques and algorithms in this paper that are widely used to improve security. In order to determine the optimal technique for encrypting the data, we also contrasted a few important observation methods and algorithms. Considering the significant role that multimedia plays, a great deal of research is necessary. It also provides the comparison of new methods in the Big data security for the encryption of the multimedia in various way in order to protect it.

8. References And Citations

[1] K. Joshi, “A New Approach of Text Steganography Using ASCII Values.” pp. 490–493, 2018.

[2] Shivani, V. K. Yadav, and S. Batham. 2015. "A Novel Approach of Bulk Data Hiding using Text Steganography." Procedia Computer Science 57 (2015), 1401–1410. DOI: 10.1016/j.procs.2015.07.457.

[3] Roy, S. and Manasmita, M. 2011. "A novel approach to format-based text steganography." ACM International Conference Proceedings Series, May, 511–516. DOI: 10.1145/1947940.1948046.

[4] Srilakshmi, P., Himabindu, C., Chaitanya, N., Muralidhar, S. V., Sumanth, M. V., and Vinay, K. 2018. "Text embedding using image steganography in spatial domain." International Journal of Engineering and Technology 7, 3.6 (2018), 1. DOI: 10.14419/ijet.v7i3.6.14922.

[5]

[6] Zhang, D. and Zhong, H. 2014. "A text hiding method using multiple-base notational system with high embedding capacity." In Proceedings of the 2014 7th International Congress on Image and Signal Processing (CISP 2014), 622–627. DOI: 10.1109/CISP.2014.7003854.

[7] Mstafa, R. J., and Member, I. S. 2016. "A DCT-based Robust Video Steganographic Method Using BCH Error Correcting Codes."

[8] G., D., L. L. Y. Zhang, M. Zhang, and X. Yang. 2017. "Title: Video Steganography in the Compressed Area." Table of Contents, January.

[9] Sudeepa, K. B., Raju, K., Ranjan Kumar, H. S., and Aithal, G. 2016. "A New Approach for Video Steganography Based on Randomization and Parallelization." Physics Procedia 78 (2016), 483–490. DOI: 10.1016/j.procs.2016.02.092.

[10] Bandyopadhyay, P. K. 2018. "Various Methods of Video Steganography." Souma Pal. October.

[11] Khosla, S., and Kaur, P. 2014. "Secure Data Hiding Technique using Video Steganography and Watermarking." International Journal of Computer Applications 95, 20 (2014), 7–12. DOI: 10.5120/16708-6861.

[12] D. Volkhonskiy, I. Nazarov, and E. Burnaev, 2018, “Steganographic Generative Adversarial Networks,” no. June.

[13] Shadi Aljawarneh, Muneer Bani Yassein, Weam Adel Talafha, 2017, “A multi-threaded programming approach for multimedia big data: encryption system” in Multimedia Tools and Applications, pp 120.

[14] Rehman, M. H. u., Liew, C. S., Abbas, A., Jayaraman, P. P., Wah, T. Y., and Khan, S. U. 2016. "Big Data Reduction Methods: A Survey." In Data Science and Engineering, December 2016, Volume 1, Issue 4, 265–284.

[15] Shah, J., and Saxena, V. 2011. "Performance Study on Image Encryption Schemes." In IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No. 1, July.

[16] Jolly Shah, Dr. Vikas Saxena, 2011, ”Video Encryption: A Survey”, in IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March.

[17] Rashmi A. Gandhi, Atul M. Gosai, 2015,”A Study on Current Scenario of Audio Encryption”, in International Journal of Computer Applications (0975-8887) Volume 116 No.7, April.

[18] Wenwu Zhu, Peng Cui, ZhiWang, Gang Hua, ”Multimedia Big Data Computing”, in IEEE Computer Society.

[19] Claudio A. Ardagna, Ernesto Damiani, 2014, ”Business Intelligence meets BigData: An Overview on Security and Privacy”, in NSF Workshop on Big Data Security and Privacy. TX, USA, September.