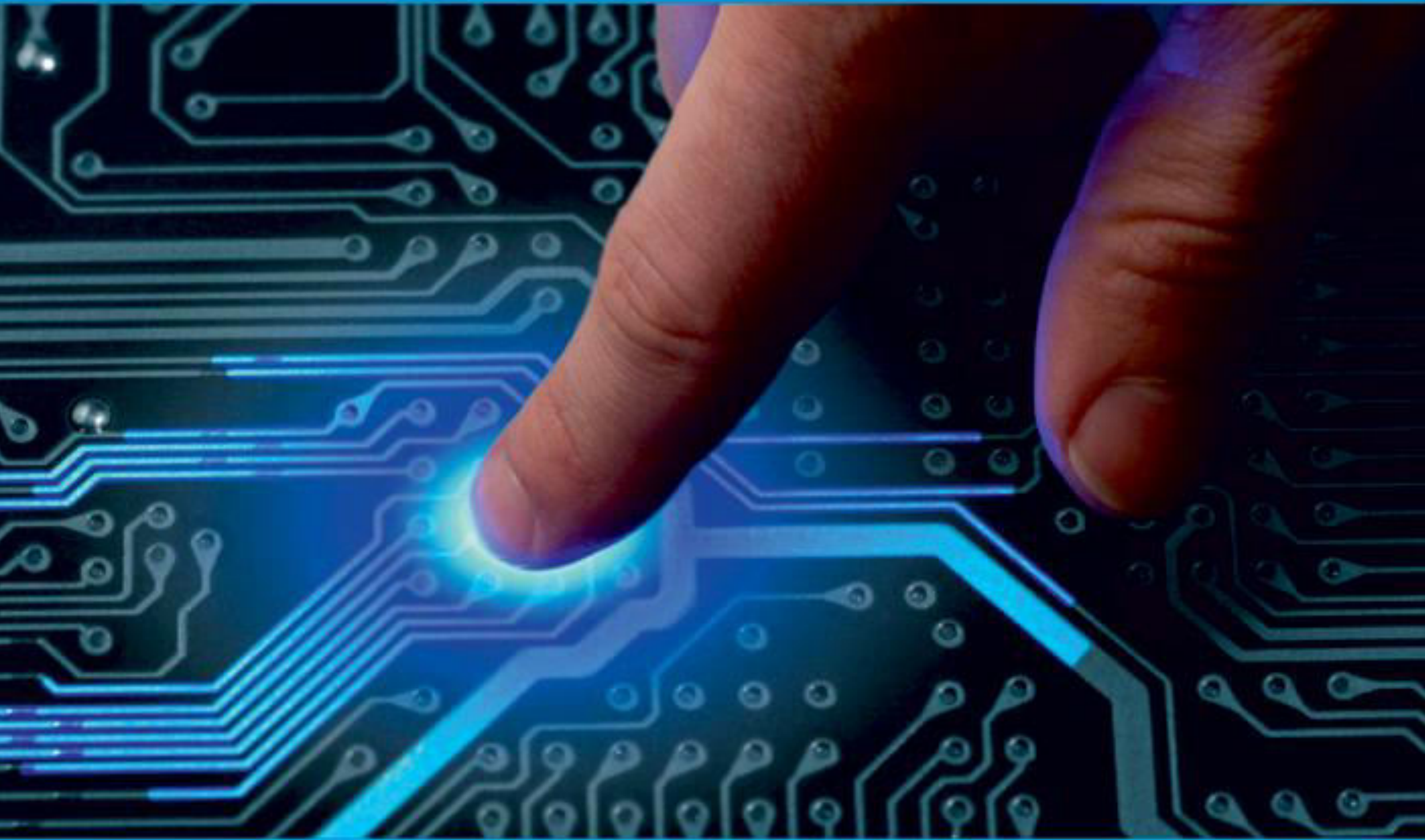
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 4, April 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004101|

An Overview of Steganography Techniques

Nagesh Jadhav[1], Nikhil Soni[2] and Prof.Swati Sah[3]

Student, School of Engineering, Ajeenkya D Y Patil University, Pune, India[1]

Student, School of Engineering, Ajeenkya D Y Patil University, Pune, India[2]

Professor, School of Engineering, Ajeenkya D Y Patil University, Pune, India[3]

ABSTRACT: Steganography is the craft of concealing the way that correspondence is occurring, by concealing data in other data. Various transporter record configurations can be utilised, yet advanced pictures are the most famous in light of their recurrence on the web. For concealing restricted intel in pictures, there exists an enormous assortment of steganography methods some are more mind boggling than others and every one of them have particular solid and flimsy spots. Various applications might require outright imperceptibility of the privileged data, while others require an enormous mystery message to be covered up. This undertaking report expects to give an outline of picture steganography, its utilizations and procedures. It additionally endeavours to distinguish the prerequisites of a decent steganography calculation and momentarily ponders which steganographic procedures are more reasonable for which applications.

KEYWORDS: Steganography, Cryptography, encryption, decryption, ciphertext, encoding

I. INTRODUCTION

One reason that interlopers can be fruitful is that a large portion of the data they get from a framework is in a structure that they can peruse and understand. Interlopers might uncover the data to other people, alter it to distort an individual or association, or use it to send off an assault. One answer for this issue is, using steganography. Steganography is a strategy of concealing data in computerised media. Rather than cryptography, it isn't to hold others back from knowing the concealed data however it is to hold others back from feeling that the data even exists. Steganography becomes more significant as more individuals join the internet transformation. Steganography is the specialty of disguising data in manners that forestalls the discovery of stowed away messages. Steganography incorporates a variety of mystery specialised strategies that stow away the message from being seen or found. Because of advances in ICT, a large portion of data is kept electronically. Thus, the security of data has turned into a central issue. Other than cryptography, steganography can be utilised to get data. In cryptography, the message or scrambled message is inserted in a computerised share prior to going through the organisation, in this way the presence of the message is obscure. Other than concealing information for secrecy, this methodology of data stowing away can be reached out to copyright security for computerised media: sound, video and pictures.

The developing prospects of current interchanges need the exceptional method for security particularly on PC organisation. The organisation's security is turning out to be more significant as the quantity of information being traded on the web increases. In this way, the secrecy and information trustworthiness are expected to safeguard against unapproved access and use. This has brought about a touchy development of the field of data stowing away. Data stowing away is an arising research region, which envelops applications like copyright assurance for computerised media, watermarking, fingerprinting, and steganography. In watermarking applications, the message contains data, for example, proprietor recognizable proof and a computerised time stamp, which normally applied for copyright security. Finger impression, the proprietor of the informational index instals a chronic number that remarkably recognizes the client of the informational index. This adds to copyright data to make it conceivable to follow any unapproved utilisation of the informational collection back to the client. Steganography stows away the discharge message inside the host informational index and presence subtle and is to be dependably conveyed to a collector. The host informational index is deliberately debased, however in a clandestine way, intended to be imperceptible to a data examination.

II. PROBLEM STATEMENT

The previous consists of etymological or language types of stowed away composition. The latter, for example, undetectable ink, attempts to conceal messages truly. One inconvenience of etymological steganography is that clients should prepare themselves to have decent information on linguistry. As of late, everything is moving toward digitization. What's more, with the advancement of web innovation, computerised media can be communicated advantageously over the organisation. Hence, messages can be covertly conveyed by advanced media by utilising the steganography methods, and afterward be sent through the web quickly. Steganography is the specialty of concealing the way that correspondence is occurring, by concealing data in other data. A wide range of transporter document organisations can be utilised, yet advanced pictures are the most well-known due to their recurrence on the web. For concealing restricted intel in pictures, there exists an enormous assortment of steganography procedures; some are more perplexing than others and every one of them have individual solid and flimsy parts.

So, we set up this application, to make the data concealing more straightforward and easier to understand.

III. LITERATURE REVIEW

What is Steganography?

Steganography is the act of concealing private or touchy data inside something that has all the earmarks of being nothing out to the ordinary. Steganography is regularly mistaken for cryptology on the grounds that the two are comparative in the manner that the two of them are utilised to safeguard significant data. The contrast between two is that steganography includes concealing data so apparently no data is concealed by any means. Assuming an individual or people sees the item that the data is concealed within the person in question will have no clue about that there is any secret data, consequently the individual won't endeavour to unscramble the data.

What steganography basically does is exploit human discernment, human faculties are not prepared to search for records that have data within them, albeit this product is accessible and can do what is called Steganography. The most widely recognized utilisation of steganography is to conceal a document inside another record.

History of Steganography:

Over the entire course of time Steganography has been utilised to covertly impart data between individuals.

A few instances of utilisation of Steganography are previous times are:

1. During World War 2 imperceptible ink was used to compose data on bits of paper so the paper appeared to the normal individual as being clear bits of paper. Fluids, for example, milk, vinegar and natural product juices were utilised, in light of the fact that when every last one of these substances are warmed, they obscure and become apparent to the natural eye.
2. In Ancient Greece they used to choose couriers and shave their head, they would then compose a message on their head. When the message had been composed, the hair was permitted to develop back. After the hair developed back the courier was shipped off to convey the message, the beneficiary would shave off the courier's hair to see the message.
3. One more technique utilised in Greece was the place where somebody would strip wax off a tablet that was

Why This Steganography?

This procedure is picked, in light of the fact that this framework incorporates subtlety as well as un-heavenliness by any steganalysis instrument. The act of adding a watermark - - a brand name or other distinguishing information concealed in sight and sound or other substance documents - - is one normal utilisation of steganography. Watermarking is a method frequently utilised by online distributors to recognize the wellspring of media records that have been found being shared without consent.

Methodology:

User needs to run the application. The user has two-tab options – encrypt and decrypt. If the user selects encrypt, the application gives the screen to select the image file, information file and option to save the image file. If the user selects decrypt, the application gives the screen to select only the image file and ask the path where the user wants to save the secret file.

- This project has two methods – Encrypt and Decrypt.
- In encryption the secret information is hidden with any type of image file.
- Decryption is getting the secret information from an image file.

Detecting Steganography:

The craft of distinguishing Steganography is alluded to as Steganalysis. To put it simply, Steganalysis includes distinguishing the utilisation of Steganography within a document. Steganalysis doesn't manage attempting to unscramble the secret data within a document, simply finding it. There are numerous techniques that can be utilised to distinguish Steganography, for example, "Seeing the record and contrasting it with one more duplicate of the document found on the Internet (Picture document). There are typically various duplicates of pictures on the web, so you might need to search for a very long time and attempt and contrast the presume document with them. For instance, assuming you download a JPED and you presume the document is additionally a JPED and the two records look practically indistinguishable separated from the way that one is bigger than the other, it is most likely you speculate the document has stowed away data within it.

Text Steganography Categories

Text steganography is considered exceedingly difficult due to the inadequate redun-dant data in textual files compared with other digital media, such as audio, image, or video files 38. Text steganography can be generally split into three classes, as depicted in Figure 5: linguistic, format-based, and random and statistical generation

Format-Based Method:

In this form of steganography, the physical features of text symbols are used. The features are altered in such a manner that the human eye cannot sense them. For example, lines in the text are moved up and down to conceal the bits of secret data. Then, words are moved left or right, or up and down. In some cases, white spaces among the words or between the paragraphs or lines are used to hide data. In feature-based encoding, the physical features of the words are altered to conceal the information. This is reliant on symbols and languages. Numerous studies into format-based approaches enhance the capacity of text steganography by changing the physical nature of the text format. For instance, maps the secret message's binary digits with the cover text's binary digits using the American Standard Code for Information Interchange (ASCII) characters, comprising punctuation, spaces, and symbols. The secret text is initially encrypted using a one-time pad and transformed into ciphertext. Then, each character is transformed into 7-bit binary numbers

Linguistic:

This technique uses linguistic steganography for hiding secret information inside text files. a linguistic steganography technique was proposed that is a topic-aware neural-linguistic steganography method. It can generate a steganographic paragraph with a specific topic based on knowledge graphs (KGs). A KG provides data about relevant topics and content to generate coherent multi-sentence texts for better concealment. The proposed method provides the quality of the generated steganographic text and its relevance to a specific topic. The author in 58 proposed a method that focuses on addressing the inability to control the semantic expression in text steganographic generated by neural networks. The author addressed control cognitive imperceptibility as a new challenge, which the steganography models must attempt to overcome in the future. The author compared three encoder models for semantic extraction, namely the Gated Recurrent Unit (GRU) model, the Transformer model, and the Topic-Aware model.

Random and Statistical Generation:

The statistical characteristics of a language are obtained and then employed to generate cover text. The work proposed a statistical text steganography technique based on the Markov Chain model, which focuses on transition probability, one of the most significant ideas of this model. This technique developed state transition-binary sequence illustrations based on the ideas and used them to regulate the production of new texts with embedded information. This technique uses the transition probability in the steganographic text generation process. This technique also encodes the state transition-binary sequence diagram required by the receiver to obtain the information, which further improves the security of the steganography information. The results showed that this model had greater hiding capacity than previous techniques. Another technique was developed by 73 to increase the embedding capacity in the statistical steganography method by implementing a Markov Chain (MC) encoder/decoder integrated with HC for steganography of Arabic text. A lower bound and an upper bound are calculated for the stego text length, which is based on the parameters of the designed encoder/decoder.

Masking and Filtering

Information is hidden inside of an image using digital watermarks that include information such as copyright, ownership. The purpose is different from traditional steganography since it is adding an attribute to the cover image for extending the amount of information presented. Masking techniques hide information in such a way that the hidden message is more integral to the cover image than simply hiding data in the "noise" level. Masking adds redundancy to the hidden information. This makes the masking technique more suitable than LSB with lossy JPEG images. It may also help protect against some image processing such as cropping and rotating. Masking and filtering techniques hide information by marking an image and is usually restricted to 24-bit and grey-scale images. Digital watermarks include information such as copyright, ownership, or licence. While traditional steganography conceals information, watermarks extend information since it becomes an attribute of the cover image.

Algorithm and Transformations:

This technique hides data in mathematical functions that are often used in compression algorithms. The idea of this method is to hide the secret message in the data bits in the least significant coefficients. Least Significant Bit Insertion: The most common and popular method of modern-day steganography is to make use of the LSB of a picture’s pixel information. Thus, the overall image distortion is kept to a minimum while the message is spaced out over the pixels in the images. This technique works best when the image file is larger than the message file and if the image is grayscale.

How is steganography used today

In current computerised steganography, information is first scrambled or muddled in another manner and afterward embedded, utilising a unique calculation, into information that is important for a specific record arrangement, for example, a JPEG picture, sound or video document. The mystery message can be implanted into customary information records in various ways. One strategy is to conceal information in bits that address similar variety pixels rehashed in succession in a picture document. By applying the scrambled information to this excess information in some subtle manner, the outcome will be a picture record that seems indistinguishable from the first picture however that has "commotion" examples of standard, decoded information. The act of adding a watermark - - a brand name or other recognizing information concealed in interactive media or other substance documents - - is one normal utilisation of steganography. Watermarking is a method frequently utilised by online distributors to recognize the wellspring of media records that have been found being shared without consent.

While there are various purposes of steganography, including installing touchy data into document types, perhaps the most widely recognized procedure is to insert a text record into a picture record. At the point when this is done, anybody seeing the picture record ought not be ready to see a distinction between the first picture document and the scrambled document; this is achieved by putting away the message with less huge nibbles in the information record.. The picture was a photograph that I had recently sent him of a spring I had rowed to while on Yellowstone Lake the previous summer.

Current Steganography

With the growth of computing power, the internet and the development of digital signal processing (DSP), steganography has gone digital. It adds words like “mp3”, “jpeg”, "Mpeg", and "text" files to our daily vocabulary. Usually these are the number of digital technologies the community faces, namely text files, still images, audio or video and documents. Today, steganography is being studied for both legal and illegal reasons. Steganography will provide the ultimate proof of authenticity that no other security tool can provide. For example the digital watermark controls the copyright of objects that are transmitted over the web such as photos, music, movies and TV shows. Average 6.53% of the size of the garbage can available in an integrated document format used for steganography [19]. The same data encryption methods apply to XML files with steganography text. From the suspects' point of view, good strategies for hiding data must meet the objectives of security and power [19]. Unfortunately, these activities are so limited in the same way and confidential information using these methods is limited. Recently, Microsoft introduced the XML file format known as OOXML through its MS Office 2007 documents.

This format contains several XML files and other existing binary files put together to form an OOXML document. OOXML document flexibility can be used for steganographic purposes and MS Office 2007 has a built-in feature that allows us to hide text in a document. MS Office 2007 also offers the "Document Inspection" feature to delete hidden information from app-generated documents and users. To our knowledge, few authors have provided the official framework for that steganography in OOXML documents and proved that it is possible. This proves that too little work is being done with this new format, i.e. OOXML format. In [15], authors hide data in XML documents using anonymous components once an unknown relationship. They also developed an acquisition algorithm to detect its presence encrypted data using the detected method. The detailed description of their work is highlighted in the data encryption section of this thesis. We believe there are many ways to hide data in them.In both cases these comments are ignored by MS Office 2007 application and these comments are rejected when writing a text back. They also investigated whether using the base64 encoding method themselves successfully embeds a file as a comment on one of the XML files of the OOXML document and MS Office 2007 is silent. Their advanced tool is called docx-steg.py and can hide any file using this data encryption program. The data encryption program is provided in the hidden part of this thesis. Their research work also focuses on XML-based documentation related to their forensic effects. This provides further research guides as these structures allow interpreting information from XML tags to detect unauthorised fraud once14 supporting court proceedings that enable law enforcement officials to use as digital evidence against the accused. We chose this study because it provides an opportunity to study several different aspects of steganography in OOXML documents highlighting deep format limits embeds encrypted data and displays techniques that make detection difficult, too identifying and developing a fast and effective OOXML document acquisition algorithm to provide for the need for real time.

Classical steganography

In this section, we will revisit related errors in classical steganography. In 1984, Simmons introduced steganogra- phy about the prisoner's problem [23], in which Alice and Bob plan to escape from prison. Their connection - you pass a prison warden looking for anything hidden- den communication between the two and when it receives one, he will separate them and make their failure cap. Then, these two prisoners must use discretion in order to keep their communication invisible. Alice and Bod must use a path that is not visible to the guard. This private channel is known as a private channel. Figure 1 shows the security model used in cryptography [24]. This connection transfers confidential information to how it can be seen by the enemy. It uses normal unprotected communication channels. The standard steganography procedure benefits from not actions may allow for retrieval message transfer. The first is the attacker's ability knowing that there is a connection between Alice and Bob. Stego media is usually sent to Bob via open or unlocked channel. The simple fact is the attacker knows that there is an exchange between two can raise suspicions. The second action is to tacker's ability to catch and read in depth modified message content to reflect the old Private channel. The attacker can benefit from this purpose by doing step analysis and finally receives or issues a private message. Attacker can proceed by closing the hidden message so that the recipient can extract it and / or modify encrypted message to send incorrect information to the recipient

Encryption algorithm

The Asymmetric Key Cryptography method will be used for the first time to encrypt the message. The Basic DES algorithm is the selected data encryption algorithm. Now, this encrypted data can be hidden in a multimedia file. Using the fixed bit text formatting, encrypted data will be saved to the image by reducing the pixel values to the nearest zero (or predefined digit) digit. A specific number that describes the 3-D representation of the letter in the order of the cipher code can be added to this number. For every character in the message a certain change will be made to the RGB pixel values. (This change should be less than 5 for each value R, G and B) This deviation from the original value will vary for each letter of the message. This deviation depends on the specific data block (grid) selected on the reference site as well. For every byte of data one pixel will be fixed, one byte of data will be stored per pixel in the image.

The cipher sequence can be specified without the original image and the recipient will only receive the corrected image. Image elements will be embedded in the first few lines of its layout, and saved to give us information when the image has been changed or the image extension has been changed such as from bmp to jpg or gif. These structures can be used to remove encryption Therefore, in short, only a code with the correct code in the appropriate scheme will create the sent message.

Related activity:

Before giving a brief description of the work related to Text Steganography, it will be worthwhile to discuss the common terms used in this context. The encrypted message is called embedded data and the pure text / sound / image used for embedding is called cover. The resulting output after embedding is called a stego-object. It is the main goal that sending and receiving is agreed upon by the protocol / key exchange process. A number of tasks have been performed to cover text. The following is a list of the various activities that have been reported and reported so far.
Generally, Text Steganography methods can be divided into two groups:
1- Changing text format.
2- Changing the meaning of the text.
Methods based on changing the meaning of the text are limited. Some examples of
These methods are as follows:

IV. DISCUSSION

This thesis aims to explore the most common methods used by Steganography, chal the length of the sight, the state of the art of self-defence against it and the tide the state of use of Steganography in attacks. Evidence suggests that Steganography is a difficult threat to deal with. This is supported by [86], indicating that a number of computer-assisted detection methods are obfuscated malware. In addition, the attack was able to escape detection for a long time which also supports this fact [6, 7, 65]. When it comes to finding information about specific tactics used in an attack, it was challenging times to find sources that have provided extensive analysis of attacks. In finding sources to answer the third research question, there was very limited discussion about protection against stegomalware. A lot of research has been around Steganography testing. This may be due to the discovery of Steganography that requires more effort than eliminating a malware program. However, adoption and self-defence go hand in hand when the discovery can be seen as one part of self-defence. Discovery and the dispute however was chosen to be discussed separately to allow for one explicit discussion about the acquisition strategies proposed in the literature. The work and research conducted for this thesis focuses on threats to companies or organisations. This is partly because the use of Steganography appears to be higher than the most common of the most advanced attacks such as Advanced Persistent Threats (APT). Available-previously not the most dramatic effect on the impact of Steganography that you may have in the general community / individuals. This does not mean, however, that people are less likely to be victims of this threat. It is shown that most attacks start with criminal emails to steal sensitive information, something that has caused population growth during the COVID-19 epidemic [69, 99]. Recommendations for So the people who should be protecting themselves from this threat would be to disconnect from the emails sent.

in addresses they do not know, not to enable macros in documents except where they are, they are absolutely sure that they are safe and maintain their operating system and anti-virus software on time. While investigating a reported attack that recently used stegomalware, it was often the case that new attacks have been reported every time you look at various websites. This provided support in the fact that attacks involving Steganography are likely to increase, which was also the case trends reported by CUING [4]. Due to a lack of data in certain areas such as how certain contracts were used in testing, the results cannot directly determine the steganographic bandwidth of all the techniques used in it the attack presented. Lack of information did not contribute to the fulfilment of findings. Moreover, due to time constraints, certain visual techniques were not possible 41 so that each one can be thoroughly explored to provide a clearer picture of the most effective technologyniques. This would be better to see if there are any strategies, for example, ML strategies that are supported, developed or work better against certain blurring strategies. No. no matter what it was, the purpose of finding the visual aids was never to go into more detail, as the results should still be the norm.

V. CONCLUSION AND FUTURE WORK

The first research question was aimed at studying the most common methods of Steganography as their function. The study found several common strategies and showed how these strategies work. Strategies include embedding information into images, using malicious macros in documents and encrypted information on HTTP, DNS, ICMP, TCP and UDP traffic. In addition, the effectiveness of each strategy was discussed based on their invisibility, capacity / bandwidth and durability. The challenges of obtaining Steganography were investigated in a second study question. The human factor has been identified as a major debt. In addition, the study of the demon- structured challenges linked to the use of common online protocols in attacks. Acquisition strategies pointed out which shows that machine learning methods show good promise. The state of the art when it came to defending against Steganography was its purpose, the third research question. Recommendations such as keeping security systems up to date with information such as and the training of people identified. Content Threat Removal can deal with digital media steganography. In addition, getting used to traffic and limiting known private channels can help to deal with the threat of secret channels.

Finally, the status of Steganography was investigated. It was not possible to provide a rate of attack using steganographic tactics. CUING has reported an increase between 2011 and 2019 and a report from the Red Canary reports that 1400 out of 20,000 threats see that they are using some kind of shading. Evidence points to the Steganography trend to keep growing and the complexity of the attacks increased again. In conclusion, this work is important to help the online security industry and re- detectives. Provides a complete overview of the most common methods used by cy- criminals sending Steganography. This thesis may be the basis for further research in this article and moreover data to detect Steganography threat is something to be taken seriously. As a result of this book review only, there is no strategic evaluation or active review of the identification / prevention methods used in real life Therefore, future work can focus on working with the company (s) to see what strategies they use to deal with the problems of the threat of Steganography. Work can also be done to review the structure of the protocol, reviewing theoretical approaches channels during the design and research phase of the current work including this research area. A very detailed discussion about each stage of finding a computer-friendly and straightforward program strategies within each phase can be developed. Best, information about tech- 43niques distributed by software to detect malicious software should be detected to give an idea of what works in practice.

REFERENCES

[1] Cloud Security Alliance, “Security Guidance for Critical Areas of Focus in Cloud Computing V3.0,” 2011.[Online] https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf.

[2] Wayne Jansen and Timothy Grance, “Guidelines on Security and Privacy in Public Cloud Computing,” NIST Special Publication 800-144, 2011, [Online] http://csrc.nist.gov/publications/nistpubs/800- 144/SP800-144.pdf.

[3] J. Nickel, Mastering Identity and Access Management with Microsoft Azure. Birmingham, UK: Packt Publishing, 2016.

[4] A. Hietajärvi and K. Aaltonen, "The formation of a collaborative project identity in an infrastructure alliance

project", Construction Management and Economics, vol. 36, no. 1, pp. 1-21, 2017.

[5] G. Goth, "Identity management, access specs are rolling along", IEEE Internet Computing, vol. 9, no. 1,pp. 9-11, 2005.

[6] Moerland, T., “Steganography and Steganalysis”, Leiden Institute of Advanced Computing Science, www.liacs.nl/home/ tmoerl/privtech.pdf

[7] Silman, J., “Steganography and Steganalysis: An Overview”, SANS Institute, 2001

[8] Jamil, T., “Steganography: The art of hiding information is plain sight”, IEEE Potentials, 18:01, 1999

[9] Wang, H & Wang, S, “Cyber warfare: Steganography vs. Steganalysis”, Communications of the ACM, 47:10, October 2004

[10] Anderson, R.J. & Petitcolas, F.A.P., “On the limits of steganography”, IEEE Journal of selected Areas in Communications, May 1998

[11] Marvel, L.M., Boncelet Jr., C.G. & Retter, C., “Spread Spectrum Steganography”, IEEE Transactions on image processing, 8:08, 1999

[12] Dunbar, B., “Steganographic techniques and their use in an Open-Systems environment”, SANS Institute, January 2002

[13] Artz, D., “Digital Steganography: Hiding Data within Data”, IEEE Internet Computing Journal, June 2001

[14] Simmons, G., “The prisoners problem and the subliminal channel”, CRYPTO, 1983

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING