

Steganography – The art and science of hiding messages within messages

Abstract

Throughout the document will be a proposal and insight of a development project that is about Steganography, which is the concept of hiding data within data. The idea is to review through the different steganographic software and have a look at the methods used to achieve a hidden message. A development task will then be set to complete for designing and developing an image steganography program used from the information which would be identified from researching the methods. Furthermore, an overall overview will be given on why this idea, how the intended aim will be accomplished, discussion of the methodology that will be followed, resources required, plan and more.

Table of Contents

Steganography – The art and science of hiding messages within messages	0
Abstract.....	0
Table of Contents.....	1
1. Introduction	2
1.1 Literature review	2
2. Proposed Research	4
2.1 Aims and Objectives.....	4
2.2 Outcome.....	5
3. Research & Development Methods	5
3.1 Methodology	5
3.2 Reliability and validity	6
3.3 Data collection	6
3.4 Risks	6
3.4.1 Time Management:	6
3.4.2 Software Crashes:.....	7
3.4.3 Hardware Crashes:	7
4. Resources & Planning.....	7
4.1 Resources	7
4.2 Planning	7
5. Conclusion	7
6. Appendices	8
6.1 Appendix A – GANTT Chart.....	8
7. References	9

1. Introduction

The working title decided is 'Steganography – The art and science of hiding messages within messages.' The project being proposed is revolving around Steganography, the process of hiding data within data. It stems from the Greek words, 'stegos' translating to 'cover' and 'grafia' meaning writing, combining both comes to the meaning 'cover writing.' Specifically, the technical use behind this gives the power of concealing any digital data which can be texts, images, video or audio content into other digital data of which is listed [1]. This project will take the tasks of going through and reviewing the existing software and different methods that are used, with the end goal being to develop and implement a steganographic software which pertains to the concept of hiding a text message within an image. Throughout the document, thorough information will be discussed on the subject area, as well as the methodology that will be discovered for reaching the end goal with regards to the aims and objectives which will be thought of and planned. The resources and planning phase of the project will be considered though with uncertainty, as it could likely be a changing aspect. For planning, a GANTT chart will be created with the time scheduled on how the tasks will be structured and worked upon. In support of any decisions made, a comprehensive literature review will be written going over several studies which include journal articles and books to identify relevant information towards the idea being proposed.

1.1 Literature review

Numerous research papers have been found relating to the topic of choice, to give a review and further background of steganography, the information which was acquired from the research papers will be collated and discussed. Primarily, further discussion will be done on the background, the uses, limits, and the methods of steganography.

One of the journal articles found, "An introduction to steganography methods", goes through the several types of methods involved in steganography and more of which delves further into the topic [2]. Part of the introduction involves stating steganography as the concealment of information, this concept naturally relates to other techniques, one they mentioned is 'Cryptography' which is the scrambling of information by transforming it into a format that is secure. Further research states this process goes by the term 'encryption' and it is for keeping information safe from unwanted recipients [3]. This information in the journal provided insight of another closely related technique of securing information, however, stating that an encrypted message "will arouse suspicion", whereas an "invisible" message with the use of steganography would not, this gives indication that one another would be used for separate reasons. Methods held and discussed within the article is text, image, and audio steganography. In relation to the intended project, this is valuable information that delivers an understanding towards the technical aspect and approaches. Text steganography is achieved through the altering of the text formatting or the certain elements of the text which would be the characters. A statement is given that the goal of designing the coding methods of alteration are to make it "reliably decodable yet largely indiscernible to the reader", making it clear that the process of retrieving

the hidden text should be as efficient as possible while being unclear there is anything hidden to the person reading. In addition, mentioned are three coding techniques, line-shift, word-shift, and feature coding which are different approaches of text steganography. In particular, the information which will be of most use is about image steganography, as the project will be taking on this task in development. To summarise the findings for hiding a message in an image, it can be altered in noisy areas of the cover to prevent as much attention that it was modified in any way. One of the common techniques used to achieve this is called 'LSB', an abbreviation for 'least significant bit', this technique is a simple approach for embedding information in an image. It is described that the bits of the message is embedded within the "least significant bit plane" of the image, and what this achieves is a hidden message with no sort of noticeable modification on the image to the user. Also mentioned is that it is best to use a lossless compression image format with the LSB approach, as there would be no compression of pixels which the approach requires bits from, if it were to be a lossy compression type, the data that is hidden would be lost. Other techniques included are masking and filtering which is constrained to grayscale or 24-bit images, this is said to be more robust compared to LSB as another path is taken for handling the hidden message. Transformations are another technique for this but is proven to be more complex, using modifications of discrete cosine transformations, also known as DCT.

Forward onto another journal article, "An overview of image steganography", this largely goes in-depth detail of image steganography and a review on the several methods for the implementation of it [4]. One of the key information speaks of two types of steganography, fragile and robust. From the explanations, fragile steganography is where the hidden information will be destroyed upon any further modification of the file, an example can be when the file format of the image is converted or making edits. The second type is robust steganography, the hidden information would not be as simply destroyed compared to fragile, but this calls for a more difficult implementation. Another key point to mention is that there are classifications of image steganography, the approaches to implementing image steganography would fall into either 'Transform Domain' or 'Image domain', both which are also known as frequency domain technique and spatial domain technique. Any "transformation and manipulation" of an algorithm would apply to transform domain, whereas "bit insertion and noise manipulation" of an image would be image domain. Given in this article are also the advantages and disadvantages of LSB, the advantages hold that it is a simple implementation method, with many techniques using it, no suspicion to be raised by looking and high perceptual transparency. The disadvantages to this technique mainly relate to its robustness, it is sensitive to filtering and modifications such as resizing, rotation or cropping would simply get rid of the secret message due to the extra noise being added.

A third journal, "On The Limits of Steganography" [5], clarifies the uses of steganography and also makes comparisons with cryptography as well as holding discussion in other areas. One key point that struck out is the use of steganography to protect intellectual property, called digital watermarking. This can serve as copyright protection with distributing digital copies of items, as it would be possible to

have embedded hidden information holding the customer name or other details for identification in the case of the copy being leaked, essentially meaning that there is an identity tailored with it and can have the person held responsible.

A book by Stefan Katzenbeisser, "Information Hiding Techniques" [6], provides an in-depth range of information all to do with steganography, one of the chapters of the book goes through the term 'Steganalysis'. Steganalysis is the process of finding the hidden information within a digital media which goes by a 'stego-object,' and this is detected by the 'steganalyst.' The reason behind steganography is to not raise any suspicion towards a message that is hidden, so no unwanted recipients would ever know. However, with steganalysis, this would go through the process of identifying the hidden messages and in some cases have them altered or removed, which raises concerns on whether watermarking intellectual property is vulnerable to this.

In conclusion of the literature review, going forward with the information salvaged will assist in making decisions of where the project will head towards. The papers served similar information but noting different points around steganography.

2. Proposed Research

The proposed research is to assess the existing steganographic software and to also identify the different methods and reasons behind using one another for achieving a hidden message. Fitting into the field being studied, software engineering, a steganographic software will be developed where the system will have the capability of taking an image and either retrieving or embedding a hidden text message. From the literature review, an assortment of past studies exhibits the history, reasoning, and uses behind Steganography, expanding the interest of wanting to do this. In addition to the reasoning, the concept of steganography dates to 500 BC [7], and intuitively the idea of it was incorporated digitally early on, as many of the past studies are not recent and are slowly becoming outdated, giving the opportunity to research more and discover of any recent advances made in relation to this topic that past research did not contain. The literature review highlights that the digital steganography is complex but serves a simple and valuable task. The information that was read was enticing and provides an ambitious task of producing a steganographic software depending on the technique that is chosen from which were discovered.

2.1 Aims and Objectives

The overarching aim of the study will be:

- To implement a steganographic software capable of hiding and retrieving a text message within an image.

Based on the literature review and understanding of image steganography the objectives of the aim will be:

1. To identify and review digital steganography software, and methods of which are used for hiding data within data.

2. To decide and learn the approach to take for implementing an image steganography program. This will be measurable by completing further research of the techniques to suit with current knowledge. It is achievable with enough research and realistic with the resources available online, books and more which include advice from others. The time it will take will require no more than a week to analyse and study the way and how it will be accomplished.
3. To design and develop the steps in achieving a program that will take an image and allow hidden text to be entered and embedded within it, including retrieving the message from one. This objective being achieved will be determined by sufficient testing of the program. It will be achievable with the already held knowledge of programming and is realistic after deciding and learning the technique and approach that will be used for image steganography. The timeframe to complete the objective with the correct distribution of tasks will see it certainly completed during the study period.
4. To analyse the developed program functionality, this includes finding the average speed it takes to process a hidden message into an image, comparing the image quality changes before and after hiding and seeing the results in contrast to popular existing programs which offer image steganography. Knowing the objective is achieved will be determined through being capable of entering hidden text onto an image within both programs. This is achievable as there is free software to allow image steganography, it is realistic as at this point the program being developed would be completed and other programs to compare it with can be found online. The time taken to find existing image steganography software and then compare it with will be minimal, making it suitable objective.

2.2 Outcome

Towards the end of the study taken place, the expected outcome is to have discovered several steganographic software with an assessment made on them. The key outcome is to also deliver a designed and developed system of an image steganography program which performs the task of seeing or hiding a text message within an image. The current vision of how it will work seeks to take an image from a file path, and then have the user enter the message they would like to embed, once that is entered the program will process and complete the task.

3. Research & Development Methods

This section will be identifying and explaining the research and development methodology to be used for gathering data and how the approach towards researching and developing will be.

3.1 Methodology

The research philosophy and development method chosen for the project in plan is on the positivist route, where research completed is independent and objective, meaning information will be based on facts. The approach will be deductive where the outcomes and conclusions will be drawn after implementation. There will be

multiple strategies involved for the capability of triangulating the results, comparison can be made from each of the chosen ones which will be, software development, document review and case study.

As part of the development aspect, a relevant methodology to make use of can be a software development life cycle, also abbreviated as SDLC. Choosing an SDLC will provide a process to follow through for the steps of designing and developing the system within reasonable time.

The typical phases within an SDLC:

- Requirement analysis
- Planning
- Software design such as architectural design
- Software development
- Testing
- Deployment [8]

There are several types of development life cycles, one being agile which will be used due to the nature of it allowing testing and development to be a continuous iteration during the development, unlike the waterfall model which is an SDLC that holds the limitation of being difficult to make alterations from previous stages as it follows a linear sequential path, further reasons such as testing being completed at the end of development make the waterfall model not work well with larger projects [9].

3.2 Reliability and validity

The reliability of the development work completed will require comprehensive testing, during the process of development and even after to assure there is no major issues in the end. A framework can be created in which can be followed each time after making a certain amount of changes within the program to assure reliability of the quality.

The validity will come from closely following the aims and objectives made for the project proposal, it stays in line with the intentions. Development should follow the design that will be created of the steps which will be the path of how the program should be created.

3.3 Data collection

The data being intended to gather will be about the different steganographic software and methods. The data will mainly be interpreted and collected from document reviews, but self-initiative can be taken for finding the software and analysing it.

3.4 Risks

The risks which require being wary of during the project are primarily to do with crashing, both software and hardware, though it also includes issues such as poor time management.

3.4.1 Time Management:

Time management can be a risk if not completing tasks within good amount of time.

3.4.1.1 Prevention measures:

- Distribute tasks in a sensible manner with enough time to complete

3.4.2 Software Crashes:

There are possibilities where software being used might crash, the main worry would be while programming as there could be a loss of work.

3.4.2.1 Prevention measures:

- Having good quality working code, with regular tests
- After performing any major changes or additions, regularly save and backup
- Keeping software up to date, as bug fixes may have been deployed

3.4.3 Hardware Crashes:

Hardware can crash unpredictably; it can be due to reasons such as the environment it is in or because it is corrupted or damaged [10].

3.4.3.1 Prevention measures:

- Hardware check-up, performing maintenance checks, clearing collecting dust and other issues
- Appropriate environment, preventing chances of overheating and dust collection
- Upgrade hardware if cannot handle applications

4. Resources & Planning

4.1 Resources

Regarding the proposed project, the resources required will only need access to a suitable computer with internet and appropriate development software. Access to these resources allow for researching and development when coming to it. Additional resources needed may be relevant books to assist in development, and further information about steganography.

4.2 Planning

A GANTT chart has been constructed to display the distribution of tasks and time frame each one should be completed by. Most tasks will take no longer than 3 weeks, whereas the development of the steganography image program will be given the most time as it will require thorough testing throughout and avoiding errors. Some tasks will require being completed first before starting as indicated by the predecessor value. This plan is possibly set to change or have additions while working on it, though thought through are the assured tasks which to be completed.

[See 6.1 appendix A for image GANTT chart](#)

5. Conclusion

To conclude the document, the proposed idea is to review steganographic software, along with the different methods. Development of an image steganography program will also be done, containing the ability to retrieve and insert a hidden message within a picture. Multiple journal articles were looked at to collate information

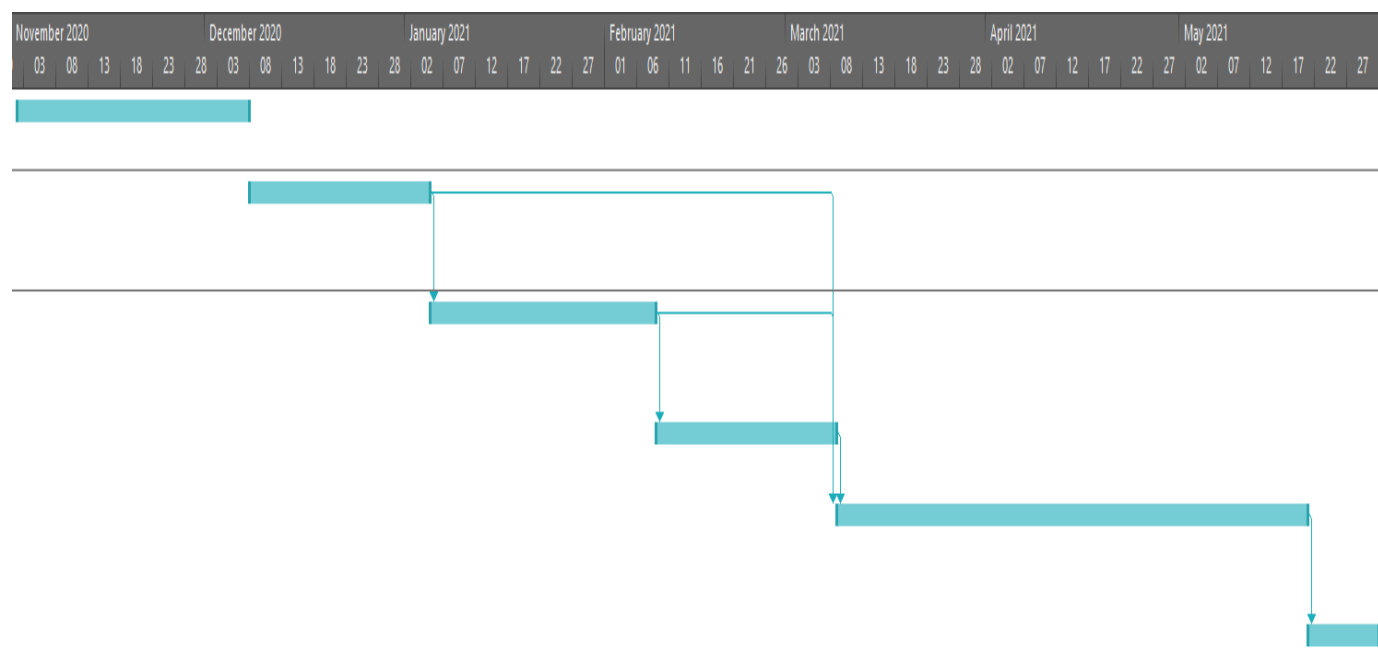
together and discuss. The aims and objectives of the project were looked at and a plan was formed with a GANTT chart giving time estimates on the distribution of tasks that should be completed for the project. Moreover, the methodology of research and development was declared to strategize how the project will manage and be completed.

6. Appendices

6.1 Appendix A – GANTT Chart

GANTT chart plan beginning from November, ending on the 31st of May.

	i	Task Mode ▾	Task Name ▾	Duration ▾	Start ▾	Finish ▾	Predecessors ▾
1		★	Find & assess digital steganography software	26 days	Mon 02/11/20	Mon 07/12/20	
2		★	Research methods used to accomplish digital steganography	20 days	Tue 08/12/20	Mon 04/01/21	
3		★	Decide and learn a method for the implementation of an image steganography program	25 days	Tue 05/01/21	Mon 08/02/21	2
4		★	Create a plan on the steps of how the program will function	20 days	Tue 09/02/21	Mon 08/03/21	3
5		★	Develop image steganography program based on plan and method	53 days	Tue 09/03/21	Thu 20/05/21	2,3,4
6		★	Analyse the functionality of the developed system and other image steganographic programs	7 days	Fri 21/05/21	Mon 31/05/21	5



7. References

- [1] M. Rouse, "Steganography," SearchSecurity, 13 December 2018. [Online]. Available: <https://searchsecurity.techtarget.com/definition/steganography>. [Accessed 28 April 2020].
- [2] M. Nosrati, R. Karimi and M. Hariri, "An introduction to steganography methods," *World Applied Programming*, vol. 1, no. 3, pp. 191-194, 2011.
- [3] C. Per, "Cryptography Definition," Techterms, 15 July 2015. [Online]. Available: <https://techterms.com/definition/cryptography>. [Accessed 28 April 2020].
- [4] R. Poornima and R. I. J, "An overview of digital image steganography," *International Journal of Computer Science & Engineering Survey (IJCSES)*, vol. 4, no. 1, pp. 23-30, 2013.
- [5] R. A. J and F. P. A.P., "On The Limits of Steganography," *IEEE Journal of Selected Areas in Communications*, vol. 16, no. 4, pp. 474-481, 1998.
- [6] S. Katzenbeisser and F. A. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, 2000.
- [7] B. Dickson, "What is steganography? A complete guide to the ancient art of concealing messages," Portswigger, 06 February 2020. [Online]. Available: <https://portswigger.net/daily-swig/what-is-steganography-a-complete-guide-to-the-ancient-art-of-concealing-messages>. [Accessed 28 April 2020].
- [8] "What is SDLC," Stackify, 8 April 2020. [Online]. Available: <https://stackify.com/what-is-sdlc/>. [Accessed 28 April 2020].
- [9] "Waterfall Vs. Agile: Must Know Differences," Guru99, 18 March 2020. [Online]. Available: <https://www.guru99.com/waterfall-vs-agile.html>. [Accessed 28 April 2020].
- [10] A. Cocilova, "Anatomy of a PC crash: 7 scenarios, and how to avoid them," PCWorld, 22 January 2013 . [Online]. Available: <https://www.pcworld.com/article/2023665/anatomy-of-a-pc-crash-7-scenarios-and-how-to-avoid-them.html>. [Accessed 28 April 2020].