

StegChat, and will be evaluated against both statistical and human analysis.

This project is a research-based project, with an emphasis on the algorithm and design rather than the implementation of the program.

1.1 Motivation

Steganography, as an area, has been the focus of less research in recent years with more time spent on improving cryptography. As will be discussed in the next chapter, this should not be the case as steganography can be a very powerful tool, especially if used in conjunction with cryptography. Text steganography is now the least popular form of steganography among researchers and commercial products. There are a number of possible reasons for this, including a much lower capacity to hold data when compared to mediums such as images or audio, but there are some reasons why it should gain more attention.

Firstly, in many parts of the world text is still the primary form of communication. This can be either through the printed word, simple messaging service (SMS) on mobile phones or email. Computers and access to the internet may also be very limited, so a steganographic technique must be suitable to be performed by hand if required (trying to insert data into the bits of an image without a computer is near enough impossible).

Text is a medium that is universally applicable. Nobody has access to digital images or audio who does not also have some access to some form of text, whether it be the printed word or digital text in the form of emails or SMS messages. Finally, text steganography, especially of the semantic form, is much more resistant to external forces than other mediums. Compression algorithms can have devastating effects of data hidden in images or audio, but you cannot compress text without losing the meaning. There is also a larger amount of variation in text compared to the other forms, making it more resistant to Steganalysis. For example, editing every k th bit in an image can be easily found by comparing it with the surrounding bits, if it is dramatically different then it will be noticed. With text, however, variations in language, even within different regions of a country with one unified language, can help to disguise hidden data, for example through synonym substitution.

1.2 Aims

The primary aim of this project is to produce a lightweight and robust steganography technique, and then to produce a demonstration application to test the algorithm in a real-life situation. The algorithm should also be simple enough that an end user can understand it without any specific knowledge. To elaborate on these points