Lightweight

The algorithm should be lightweight in two senses. The algorithm should be simple enough that it can be performed by a human using a printed dictionary. This is primarily so that the algorithm can be performed where there is no access to a computer or the internet. It should also be lightweight enough that it can be run in almost real time for the chat-based application, and can be used on a device with little computing power.

Robust

The algorithm needs to be robust against two forms of steganalysis, automatic and human. The output of the algorithm should pass any statistical steganalysis method, including frequency analysis. The text that is outputted should also be readable by humans, and should not arise any suspicion due to incorrect word replacements or a loss of meaning.

The primary aim of the project is for more robustness against steganalysis rather than a higher bitrate. The desired bitrate is around 1 bit for every 10 words, which is in line with the approaches described in section 2.3.

1.3 The Report

This report will consist of four further parts. First, there will be an explanation of steganography with a focus on text steganography, and a literature review of current research into synonym-based steganography (and similar methods). Next will be a full explanation of the design of the algorithm, and a brief design of the StegChat test application. Thirdly, there will be an section on the implementation of the algorithm and test program, including the problems implementing the algorithm and a description of the structure and operation of the program. This will be followed by an evaluation of the algorithm, and finally a conclusion.

1.4 Glossary

Synset - A set of synonyms.
Obfuscation - The act of hiding data within a covertext.
Deobfuscation - The act of extracting data hidden in a covertext.