

Matthew Tilton

The Algorithm

Step 1. Check hash of corrupted.docx

Step 2. Convert file into hex

Step 3. For each pair of set of headers and footers find all the occurrences of them in the file

Step 4. For every combination of header and footer locations take a subset of the original file.

Step 5. Check to see if the hash of the new file is in the list of correct hashes. If it is append correct on the end of the name so it is easy to notice.

Step 6. Write the file.

Step 7. After all the files are written go back and check all the potential docx files manually since I don't have a hash for them.

Usage

To use this program modify the filepath variable to point to the location of the file on your system. Then run filecarving.py in python 2.7. Once it's run check all of the files it generated and delete the files that won't open.