Code Cracking: Who Murdered the Somerton Man?

Yami Li and Yifan Ma, Supervised by Prof. Derek Abbott and Dr. Matthew Berryman

School of Electrical &Electronic Engineering, The University Of Adelaide, Adelaide, SA

Background:

On the morning of 1st December 1948 , the corpse of an unidentified man was found on the Somerton Beach. The man is always referred to as “The Somerton Man”. As a paper scratch was found in the fob pocket of the dead man’s trousers with the Persian phrase Tamam Shud (i.e. “The End”) printed on, the case is also known as the Tamam Shud case. For 65 years the Tamam Shud case has endured one of Australia’s most mysterious and peculiar case.

Figure1. The Somerton Man

Figure2. The Somerton Man

How is the case related to the project:

After Police’s investigation, the aforementioned paper scratch was proved to be torn from the final page of a book. Around six months later the book was found, with the handwriting of the mysterious code on its back. Thought myriad people devoured to crack the code, it has remained mysterious for over 60 years. Hence one aspect of the project was analyzing the code to drag it closer to the truth.

Another important clue is the hair of the Somerton Man. It was discovered from the plaster bust made after the Somerton man's autopsy. In this project, the mass spectrometer isotope concentration data of the hair was analyzed. It is expected to reveal the environment the Somerton Man had lived in before his demise.

Figure3. The Paper Scratch

Figure4. Mysterious Code on the Book

Figure5. The Plaster Bust

Scope of this Project

Based on the two key clues in the Tamam Shud case, this project was divided into two parts to work on each clue separately.

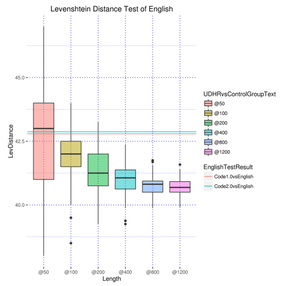
Firstly, based on the previous project group's conclusion that the code was initialisms (first letter of each word), the code has been critically analyzed. It was expected to find out which language the code was expressing by calculating the similarities between the code and varies of texts in 9 different languages and analyzing the results statistically.

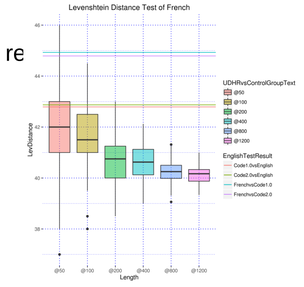
The other aspect was dealing with the mass spectrometer isotope concentration data. Data of different chemical elements' contents from the Somerton man's hair and hairs in control group were compared.

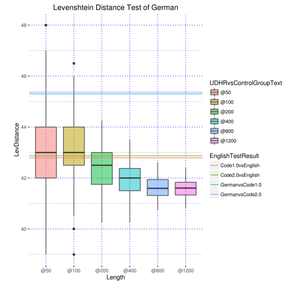
Section A: Code Analysis

Here presents the general research procedures in this section:

1. Prepared texts for comparison use. The ***Universal Declaration of Human Rights*** (UDHR), the ***War and Peace*** and their translations were selected. To meet the code's standard, all texts were transformed into initialisms using Java and *gedit* text editor.
2. The **Levenshtein Distance** and the **Simhash** algorithms were implemented and customized using *Java*.
3. Similarity check tests had been designed and carried out. Results were saved in .csv format.
4. The results had been analyzed, reshaped and selectively plotted using R.
5. According to the test results, reasonable conclusion has been drawn out.







As the figure above illustrated, the Levenshtein Distance between the code and English text is the minimum (lower distance means higher similarity). It implies that the code is more similar to English text rather than others.

2. Simhash Test

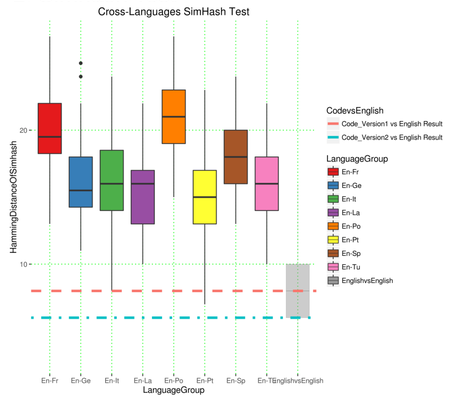


Figure6. Cross Language Simhash Test

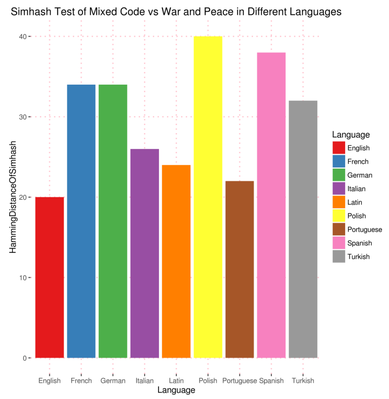


Figure7. Code vs Different Languages

Figure 6 illustrates the cross-language Simhash Test result. Figure 7 shows the Simhash test results between code and different languages. It is obvious that the Simhash algorithm can distinguish two different languages perfectly, on top of that, the result between code and English text was again the minimum one.

Conclusion: The code consists of English initialisms.

Section B: Mass Spectral Analysis

1. Pb-Content analysis

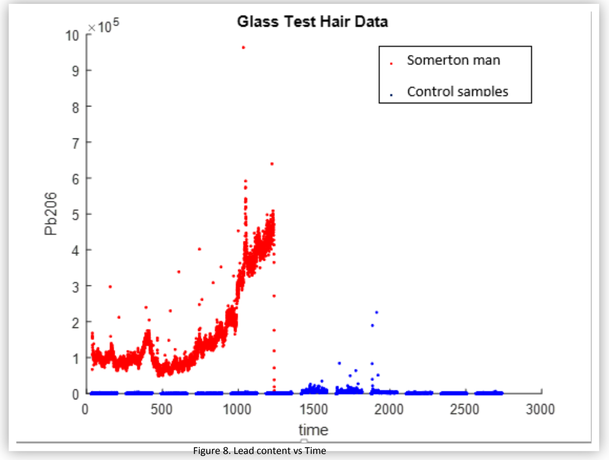


Figure 8. Lead content vs Time

This figure describes the Pb content of Somerton man’s last two weeks of life The left side of the figure is the most recent. Lead content decreases before his death

Conclusion: Lead poisoning was not the cause of his death.

2. Quartz Test

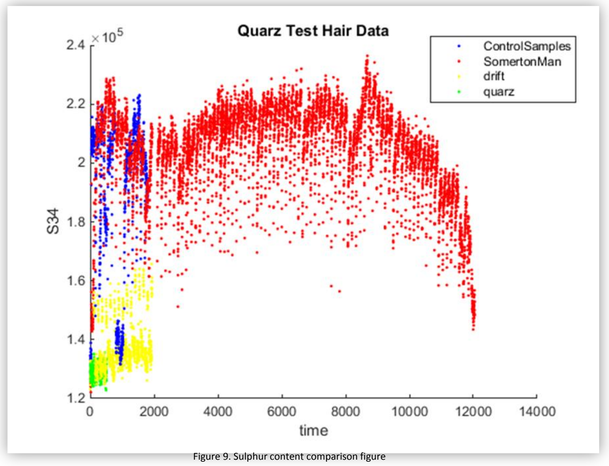


Figure 9. Sulphur content comparison figure

The relative values of sulphur are similar among Somerton man and 12 control hairs. Only two samples have significant differences and these samples come from one special person. The living environment conditions for Somerton man and six control samples are similar but different from the special one.

Acknowledgement

The authors would like to thank Prof. Derek Abbott for his kind advices and helps. Also thanks all the open-source softwares that were used in this project, they are: Java, R and Linux openSUSE.