00_Narrative

Wei Ting

5/7/2020

This document gives briefly introduction of the project, main challenges I encountered and solutions I tried. Finally the results is presented as well as some possible further works.

Project: A Prelimilary exploration of cybersecurity.

I have long been interested in the topic related to cybersecurity, especially how the orthodox theory in international relations applies to cyberspace. In this project, I explore the data from the cyber incident tracker and try to answer questions below by descriptive analysing and visualization.

Research Questions

1. Trend of cyber operation

2. What are the most common types of cyber operations ? Trend by years?

3. Countries are more likely to be Sponsors (top 10?)?

4. Frequency of cyberattacks conducted by top five countries?

5. Do super powers (US, China and Russia) more likely to sponsor cyber incidents, compared to middle power?

6. How do victims respond to the attacks?

Challenges

Collect Data

The data I use for the project is the only dataset I can find regarding cyberattacks sponsored by state level so far. I suppose that it's because cyberattack targeting states are sensitive (since I did find some data regarding general cyberattack by private sectors) and those operations are hard to track.

Cleaning- Try to seperate the "Response" column?

Since the column includes random delimiter, it can't be separated simply by " "

Sol: - I turn to excel and try to figure out what's the structure of this data. - After checking, I now know there are 68 rows containing information I want, but they may be delimited by four or five spaces. - So I broke them into five columns and made a list with same length - Make a new dataframe and refined it into a column containing words and the other containing link.

Analysis-Sequence Analysis (How to start out/ try and error)

I did not even know how to start out at the beginning! I take advantage of online resources and experience a series of try and error.

Sol: - want to know "How often does a cyber incident happen?" - First, I should find out what are top 5 state sponsors since it's not very informative to know some occasional incidents and sponsors (and don't want to make graph messy!) - After identifing the top 5, I made another column "maintype"to classify "Sponsors" column. - Then, I applied something found online to count the time gap.

Further works

Though these are relative elementary analysis, I am glad that I can use these technique after learning computational tools, and to explore more information to support my research.

For further research, I hope to incorporate some other variables such as military power composite or telecommunication data from ITU to explore the relationship between traditional state power and cyberpower in the Informational age. Given that data regarding this topic are either too sensitive to disclose, or only contain information retrieved from news, I will try other tunnels, ex. to web-scraping by myself, in order to gather more informative and comperhensive data.