On gaps in enterprise cyber attack reporting

DECISION: accept

AUTHORS: Abu Hajizada and Tyler Moore

Summary of Reviews

- Review 1: 3 (5)
- Review 2: 1 (4)
- Review 3: 1 (4)

Reviews

Review 1

TOTAL SCORE: 3 (strong accept)
Overall evaluation: 3 (strong accept)
Reviewer's confidence: 5 (expert)

The author of this paper has created a dataset on reports on cybersecurity incidents in the US by analysing SEC, HHS and Hackmageddon datasets and combining the results. Furthermore, the dataset has been published alongside the paper.

The paper clearly describes how the data was gathered, how it has been analysed and any further necessary steps to make this usable. The first analysis of this dataset already shows some very interesting results that provide enough grounds for discussion.

Review 2

TOTAL SCORE: 1 (weak accept)
Overall evaluation: 1 (weak accept)
Reviewer's confidence: 4 (high)

This paper collects breach data from multiple sources in order to understand whether companies are complying with breach reporting requirements, specifically, US SEC and HHS disclosure regulations.

The datasets are clear, and appropriate. However, identification and analysis of breach information in SEC filings is not entirely clear. Presumably the authors will have an opportunity to develop a rigorous and transparent approach.

Review 3

TOTAL SCORE: 1 (weak accept)
Overall evaluation: 1 (weak accept)
Reviewer's confidence: 4 (high)

The paper aims to analyze the prevalence of cyberattack reporting using public sources of information such as breach-notification laws, regulatory filings, and media reports. The paper quantifies the number and types of cyberattacks reported in each data source and compares these findings to crowdsourced reports of cyber incidents appearing in media outlets, revealing significant gaps in coverage.

The methodology adopted in the paper is purely empirical and quantitative in nature. However, the process was not fully automated and mostly required manual investigation by the authors, leading to results/quantities that must be taken with caution. The authors did not reflect on the limitations of the method used to search SEC filings, which could be helpful for future research.

The paper seems to be mistakenly using the concept of "cyberattack" to refer to cyber incidents. Regulatory efforts encourage the disclosure of incidents and not actual attacks, given the potential volume of attacks. In this sense, the paper does not clearly define what the SEC filings are meant to report and why the search with the term "cyber" would lead to a comprehensive set of incidents.

Some parts of the methodology rely on manual inspection of the paragraphs before and after the mention of the term "cyber" to determine if it describes an attack, introducing subjectivity and potential bias. The paper does not mention any type of coding or how the types of attacks were mapped from the text. It would also have been useful to see if the filings helped to estimate the severity of the incident. The collected information does not provide any insights about the severity or impact of the identified attacks, limiting the usefulness of the results in assessing the overall risk of cyberattacks to the companies.

Despite these limitations, the paper provides a useful starting point for future research on cyber incident reporting practices. By identifying the types of attacks mentioned in SEC filings and the language used to describe them, which can help inform the development of more effective disclosure policies and regulatory frameworks, and also be leveraged for quantitative research studies.