Applying Neutralisation Theory To Better Understand Ransomware Offenders

DECISION: accept

AUTHORS: Lena Yuryna Connolly, Hervé Borrion, Budi Arief and Sanaa Kaddoura

Summary of Reviews

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

Reviews

Review 1

TOTAL SCORE: 2 (accept)
Overall evaluation: 2 (accept)
Reviewer's confidence: 4 (high)

This paper leverages crime theory of neutralization in order to explain rationales for developing tools to facilitate, commit, and justify, ransomware attacks. In order to address these issues, the authors are reviewing past interviews conducted by other organizations.

This is a very interesting paper that seeks to apply an existing criminology theory to ransomware attackers. The sample size is an obvious limitation, but the authors describe a rigorous, structured approach to the interview analysis (i.e. developing a thematic codebook, and decomposing interviews into incidents/units).

While the research effort is still in the early stages, not having any robust analysis or results, this could turn into a wonderfully interesting paper.

Review 2

TOTAL SCORE: -2 (reject)
Overall evaluation: -2 (reject)
Reviewer's confidence: 4 (high)

This paper applies neutralization theory to understand how ransomware operators rationalize their activities. To do so, the authors analyze nine documents in which a ransomware operator was interviewed by someone at a cybersecurity company, who then published the discussion in a blog or a white paper.

The paper is well-written, and the study is intriguing, but the manuscript is incomplete and cannot be published as is. The authors provide a thorough introduction, a description of the theoretical framework, and an explanation of the method used. However, the preliminary results are presented briefly, showing how three quotes were interpreted by a coder (about half a paragraph long). Although preliminary studies can sometimes be published, more results must be presented than what is currently available in this manuscript. Therefore, I cannot recommend this paper for publication. It falls short, and as a reviewer, I cannot appreciate the relevance of the findings.

To improve the manuscript, the author may want to consider the following additional comments:

- The authors should exercise caution when stating that they study RaaS operators: the individuals interviewed may be affiliates using RaaS operators rather than the operators themselves.

- On the first page, in the third paragraph of the second column, there is a problem with the interpretation of the studies. The author mentions two studies [15][16] and then quotes [12], which states that the main limitation of these studies is that they rely upon college or university samples. The issue with this statement is unclear. If [12] means that the studied participants are volunteers from these institutions, neither [15] (which studies those behind booting services) nor [16] (which presents a CrimeBot capturing forum data) involve college or university participants. If the authors want to point out that the problem is data collection by academics, they need to develop their point further, as it is not clear why this is a problem (rather, it seems like a positive aspect).

- The author mentions that the study uses data directly collected from alleged offenders, yet their use of the data is secondary. This statement is misleading, as it implies that the authors collected the data themselves, which is not the case.

Review 3

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

The paper seeks to examine the relevance of neutralization theory to understand ransomware offenders. The theoretical background on neutralization is well understood and presented by the author(s), and the methodology used to verify whether it applies to this particular form of online offending is sound. However, I feel the paper is not fully developed yet and the research should be finalized before it is published.

The theoretical background is well-known and the idea to apply it to online offending is sound. It has been done for other specific forms of cybercrime and it makes sense to see whether it can help us better understand ransomware offending. However, it does not take into account the specifics of ransomware operations, such as the fact that many of them operate as firms, include a nationalistic dimension (and therefore see victims as adversaries deserving what happens to them), or that their attacks can harm the life of victims (such as when they hit hospitals for example). This is not really reflected in the paper.

Also, the coding process, as described, is sound, so I am wondering why the full preliminary results were not presented in the table, even if it is understood that a second coder is being brought in. The six neutralization techniques identified are listed, but no details are provided, which is a bit disappointing. So either the author(s) does not have confidence in her/his findings, or they are too preliminary to be presented at a conference such as WACCO.