[2277] POSTCOG: A tool for interdisciplinary research into underground forums at scale

PRELIMINARY DECISION: accept

Summary of Reviews

- Review 1: 1 (2)
- Review 2: 2 (5)
- Review 3: 2 (5)
- Review 4: 0 (4)

Reviews

Review 1

TOTAL SCORE: 1

Overall evaluation: 1 (weak accept)

Reviewer's confidence: 2 (low)

The authors present a web frontend, POSTCOG, which enables access to a large collection of underground forum datasets. The aim is to support an interdisciplinary setting, in which some researchers may not be familiar with the scripting environments that are typically used to explore these datasets.

The review of previous work on underground forum analysis is solid, and highlights a few main areas of interest: key actors, underground economies, social networks, trending discussion topics, which are accessed by a variety of content and metadata searches.

POSTCOG acts as an interface to a JSON export from a PostgreSQL database. It features keyword search, metadata filtering, a thread view for posts, and some automated labels for 'type' and 'intent' as well as crime type (which methods are drawn from previous work), with modularity to enable other labels. Many of the features are general to database GUIs already available, but the thread view, modular support for labels, and the integrated ability to report incorrect labels shows some customisation that would be useful for interdisciplinary forum analyses.

I was disappointed that the interface lacked features supporting more of the analysis methods identified in the literature review. For example, the system could be improved with modules supporting some forms of social network analysis, or extracting/identifying indicators of economic activity. The ability to extract random or stratified samples of posts was also identified as generally useful in the literature review, but not mentioned as implemented in the frontend.

As the contribution is aimed at improving the access of users, it undertakes usability testing with four users, using a 'think aloud' process in two sessions. The description shows ongoing improvement of the interface in terms of intuitive use, which is encouraging, but it is not clear if issues raised during the second session have now been addressed. One issue from the first session -- scalability -- is very important, and a discussion would be welcome on how the system balances the need to support complex queries on large datasets with the need to retain a responsive interface that facilitates exploration.

My view is that POSTCOG certainly has the potential to fill a niche in database GUI design that would be useful for the interdisciplinary projects it targets. However, the tool appears to still be in early stages of development and may need more refinement.

Review 2

TOTAL SCORE: 2

Overall evaluation: 2 (accept)

Reviewer's confidence: 5 (expert)

This is an interesting paper that describes a tool that should facilitate access to data for cybercrime researchers from a variety of disciplines and with limited technical proficiency to collect the data themselves on hacker forums. The need for such a tool is obvious, as lack of data is often a problem, and this paper enhances the existing literature by showing how a new web-facing interface that simplifies considerably the querying process has been built and can broaden access to this data. THe fact that the data can be exported in a universal format (csv) to perform a broad range of analyses is also a major plus. I am looking forward to testing this tool myself.

Among the issues that deserve attention for future publication:

1. the references are very Cambridge-centric and should probably be made a bit more inclusive of the work done by others on cybercrime forums (such as Décary-Hétu, Holt, Christin, Lusthaus, Dupont, Burruss, etc.). Some researchers who work on forums choose not to name them in their paper and therefore did not show up in the literature search. This should probably be accounted for.

2. It would be interesting to name the forums that are available (especially in Table 1), because otherwise, interpreting the relevance of the data used or selection of papers is impossible.

3. The testing involved four users: maybe it is a bit presumptuous to state that such a small sample was "representative of the current user base and the potential users" (p. 7). The testing was done very thoroughly it seems, so no need for hyperbolic statements like this one to still show it is a very valuable piece of work.

4. The references on the use of SNA to study cybercrime forums (#22 to 25) do not seem to include any criminology paper, while this discipline has made some interesting theory-driven contributions to the conversation. I would suggest the authors cite a couple.

Overall an enjoyable read and an interesting paper. Would be good to have some usage statistics about the database: how many researchers have been granted access to it outside of the host institution and how many papers were published in how many disciplines thanks to this database? In other words, are cybercrime researchers taking advantage of it?

Review 3

TOTAL SCORE: 2

Overall evaluation: 2 (accept)

Reviewer's confidence: 5 (expert)

This paper presents an analysis tool (PostCog) to help mine insights from underground forum measurements based on an extensive literature review and is tested on an expert user population using AnonymousDB.

Strengths

+ Very welcome contribution to the research community analysing underground forums.

+ Well-structured and readable paper, that adds - an often neglected - literature review and usability study to tool development.

Weaknesses

+ Application of the tool ‘limited’ to underground forums, whereas work on the underground economy is broader then this - e.g., anonymous online markets, or Telegram channels.

Comments for the authors

I really appreciated the thoughtful work that went in to this paper. The paper is structured well, and its contributions are clear. Instead many other ‘tooling-papers’, the authors try not to help improve the state-of-the-art in one niche area by a little for only a few, but apply the state-of-the-art in a tool for the many. In essence, the authors worked on a tool that is intrinsically valuable for all in the research community, but no one bothers to put resources in. My only suggestion would be to the authors to elaborate on the potential application of their tool on other parts of the underground economy - i.e., markets, Telegram channels.

Review 4

TOTAL SCORE: 0

Overall evaluation: 0 (borderline paper)

Reviewer's confidence: 4 (high)

Summary

This paper presents PostCog, a web-based tool that can be used by non-technical researchers to obtain and analyze underground forum data. The authors perform an extensive survey of prior works on underground forums to determine the functionality of the tool.

Comments

First, this paper gives a great survey of the different types of underground forum research that have been done in the past. Two small nit-picks here, though:

1) the categories (bolded) are more emphasized than the subsections and this makes it a little hard to read. This part of the paper is also stuck in the introduction but should be its own section.

2) Normally, I'm very pro removing lines from tables, but Table 1 is really hard to read without lines. Either the long list of citations should be broken over multiple lines or black lines/dots should be added to the table. When printed, the checkmarks are also invisible; please make them black.

Second, the user study is quite limited. The authors present PostCog as a way for non-technical researchers to access and analyze underground forum data - i.e. those who have not been able to do such research before due to technical limitations. However, the participants selected all have had experience in some way with AnonymousDB. Additionally, one 4 participants were recruited (and hand-picked). A wider user study should be carried out which includes users with no experience with Anonymous DB, and in my opinion, this is a clear bar for acceptance.