

## Peer Review

# Review of: "Quantifying Hot Topic Dynamics in Scientific Literature: An Information-Theoretical Approach"

Paweł Lula<sup>1</sup>

1. Department of Computational Systems, Cracow University of Economics, Krakow, Poland

The topic of the paper and the approach proposed by its Author are very interesting, and the method can be very useful in the analysis of huge collections of documents. The method proposed in the paper allows for the analysis of the popularity and importance of topics discussed in research publications. I highly rate the article, but at the same time, I would like to formulate a few comments that perhaps will encourage the Author to make some modifications.

- My first question is related to the problem of concept representation. It is obvious that concepts are represented as a distribution over words from a given dictionary. But how was the dictionary defined? Did the Author use one dictionary for the entire corpus (I assume that this solution was used), or was a separate dictionary used for every year? This issue has an impact on the formulas for probability calculation (equations 2 and 3). In my opinion, a short clarification and formulas for probability calculation could be added to the text,
- How is the set of all concepts for a given corpus defined? I am asking about this in the context of step 2 of the algorithm presented on page 10,
- In my opinion, a more precise description of the method of producing Figure 1 is required. I would like to propose adding a detailed description and interpretation of the minimum spanning trees created on the basis of two components of the “v” matrix and the method that was used for drawing curves between elements related to consecutive years (I can only guess that the weights of the MST for two periods were compared),
- Figure 1 does not differentiate between the seed concept (cybersecurity) and concepts related to the seed concept. If the circles for related concepts present their similarity to the seed concept, how should the circles for the seed concept be interpreted?

- How were “attraction” and “repulsion” on Figure 2 calculated?

I would like to congratulate you on your work and wish you further achievements!

## **Declarations**

**Potential competing interests:** No potential competing interests to declare.