Reviewer’s report

Title: Predicting metabolite-disease associations based on KATZ model

Version: 0 Date: 08 Aug 2019

Reviewer: Olga Lazareva

Reviewer's report:

My comments are related to the form of the paper (grammar, wording, etc) and to the method itself.

Form:

1) The quality of writing requires improvement:

a) Many sentences are too complicated, unclear and grammatically wrong.

Example:

"Thus, it is indispensable to spread computational methods which can save experimental time and fund especially, available prediction results."

"However, only depending on these methods are vulnerable in terms of cost and time."

b) There are multiple typos and incoherent spacing before/after dots or commas.

2) The meaning of acronyms is not given when the acronym was mentioned first (example: GIP kernel, DAG).

The method:

3) In the Background section authors stated:

"some relevant methods about predicting have been delivered for genomics such as gene-disease correlations [9-11], transcriptomics like circRNA-disease associations [12, 13] and proteomics such as identification of essential proteins [14-16], but the computational methods for predicting metabolite-disease associations can be counted on the fingers of a hand."

This is a very vague statement regarding competing methods. It is unclear whether those methods exist or not and if they exist, they should be mentioned explicitly.
4) Meaning and basic intuition behind KATZ framework are missing. Also, all related concepts needed for this model are not introduced.

Example:

"Generally, what need to be taken into consideration when computing the potential association between metabolite I and disease j in the known metabolite-disease associations network are the walks' number of metabolite I and disease j and the different length of different walks [22]."

Here authors write about "walks" for the first time and never explain exactly what are those walks, what are they used for and why the length of the walks matter.

"Meanwhile, there is a detail cannot be ignored that the longer walks are supposed to have lower influence than shorter."

The authors have never elaborated this statement as well even though it is important for algorithms parameters choice.

5) Section "Comparison with other methods":

Benchmarking of the results should be done with respect to the existing methods while authors just mention something about a classic random walk method and page rank that they have implemented themselves and did not give source code.

6) No available source code.

7) No information about computational complexity or run-time. This information is also missing in benchmarking with other algorithms.

**Level of interest**

Please indicate how interesting you found the manuscript:

An article of limited interest

**Quality of written English**

Please indicate the quality of language in the manuscript:

Not suitable for publication unless extensively edited
**Declaration of competing interests**

Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license ([http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/)). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal