Author’s response to reviews

Title: Analyzing hidden populations online: topic, emotion, and social network of HIV-related users in the largest Chinese online community

Authors:

Chuchu Liu (chuchuliu1993@foxmail.com)

Xin Lu (xin.lu@flowminder.org;xin_lyu@sina.com)

Version: 2 Date: 12 Oct 2017

Author’s response to reviews:

Dear Editor,

Thank you for giving us the opportunity to revise the manuscript. We appreciate the helpful feedback provided by all reviewers and editors.

We have revised the manuscript according to your suggestions and have responded to the reviewer’s comments in the attached point-by-point letter.

The main changes in the revision are as follows:

(1) We have reworked Figure 7 for better interpretation of the distribution of topics in the HIV communities.
(2) We have added a validation of the sentiment analysis result in the Supplementary Information.

We hope that we have been able to address all the suggestions raised by you and by the reviewers and believe that the reviewing process have helped in improving the paper substantially. For detailed changes in the manuscript and our replies to all the reviewers’ comments, please see “point by point response to comments by reviewers” in the following pages.
We look forward to hearing from you again.

Best regards,

Xin Lu, Chuchu Liu

Point by point response to reviewers’ comments

Reviewer #1:

The authors have produced one of the clearest and best-presented responses I have seen this year - thank you for taking the time and effort to do this properly. I only have minor comments to add.

1) I still think it would be worth presenting the p-values exactly, even if they need to be expressed in the form of 5.43x10^{-21} but I am happy to leave that decision up to the journal style.

Response: Thanks for the suggestion. To meet the journal style, we adopt the common practice of statistic writing for p with three digits.

2) Figures 3 and 7 are still relatively hard to interpret but I understand the inherent difficulty with visualising these types of data and I don’t have any alternatives to suggest. In Figure 7 in particular, I am not sure how the different shading in the inner and outer parts correspond to the legend given that the outer ring has different shading compared to the inner ring.

Response: Thanks for the suggestion. We have reworked Figure 7 for better interpretation of the topics in the HIV communities. And we have unified the shading in the inner and outer parts corresponding to the legend.

Reviewer #2:
Have the results of the sentiment analysis methods been validated? Sentiment analysis based on word counts is known for its inaccuracy.

I understand the justification for choosing the specific sentiment analysis method that you choose. What I would like to see is a validation of the sentiment analysis result. One simple way to do is to pick out, e.g., 100 random posts from your sample, ask two people manually annotate these 100 posts as the gold-standard, and compare the sentiment analysis results against this gold-standard.

Present the kappa agreement between the two people and present the precision, recall of the sentiment analysis result.

Response: Thanks for the suggestion. We have added an experiment to validate the results of the sentiment analysis in the Supplementary Information.