JRSS-SA-Nov-20-0251: "Comparing the Real-World Performance of Exponential-family Random Graph Models and Latent Order Logistic Models"

Comments to the authors from the Joint Editor

Many of my comments overlap with ones from the referees and the Associate Editor. In the revision, you should of course address all of their comments, including the ones not touched upon here.

The motivation of several of my comments is that for this journal the paper should be, as far as possible, accessible also to interested readers who are not experts in network analysis. This means, for example, explaining some things more carefully than you would in a specialist journal and avoiding or explaining network analysis jargon where necessary. When preparing the revision, please consider the text also from this point of view, even beyond my specific suggestions below.

1. As noted also by the reviewers, the text is currently a mess and in need of thorough copyediting. This applies to the main text (including careless editing, ungrammatical expressions, bad punctuation, undefined acronyms etc.), tables and figures, and bibliography. Some of my comments mention instances of this, but they are not meant to be comprehensive.

2. When revising the paper, please aim to do this without increasing its length. The revised paper should be no longer than 30 pages in the current formatting.

3. Sections 2 and 3 are currently somewhat fragmentary. I would suggest that you combine them into one section on the definitions and theory of these models and make it more self-contained and coherent. For example, this could include the following:
   (a) Describe both LOLOG models and ERGMs in roughly the same way, in matching notation and covering the same topics (with perhaps some more detail on LOLOG, but not taking ERGM as known to the extent you do now). This will make it easier to follow the discussion of the comparisons between them.
   (b) In this discussion, clearly separate model specification, interpretation, estimation, explicit comparisons between the two models, and other such topics.
   (c) Please give some more explanation of the logic of LOLOG. At first reading it seems confusing to see a time variable t and references to time ordering, given that most network data have no such observed ordering. If I understand this correctly, this is a hypothetical (latent) device which is used define the model, and the distribution of the observed network is regarded as an average over all the possible orderings. This is clear enough on second reading of the paper, but saying it even more clearly would be useful, especially since the paper would serve as an introduction to LOLOG to many readers.
   (d) Give some more explanation (and examples) of what the graph statistics g may be. Are all the covariates in your examples (and more generally) of this kind?
   (e) Say something about the interpretation of the parameters theta in both models. Without this, it is harder to understand what to make of the examples that come later, and the comparisons between them.
   (f) p 5: "… the expectations and standard errors of the MLE and MOM parameter distributions do not exist either" – so where do the estimated standard errors in your results come from?
   (g) I would suggest that you explain the goodness of fit method that you use (now on p 11) already in this section, giving also some more explanation of what it does.

4. The formatting and labelling of the tables and figures should be improved:
   (a) Make the captions more self-contained and self-explanatory.
   (b) Explain what the numbers shown in the tables are.
   (c) Put the explanation of the significance stars in a footnote of a table rather than caption.
   (d) Remove the underscores in "University_2005" etc.
   (e) In the figures, overrule the ggplot2 defaults on labels of the plots and legend texts.

5. Please give better explanation of the details of the example in Section 5.
   (a) Please give a much clearer explanation of what the results actually mean, in terms of substantive research questions about interactions of people in offices. How are the estimated coefficients interpreted, for LOLOG and ERGM? How do these interpretations compare with each other? You now start to say something about this in Section 5.3, but this is still obscure and a little too late.
   (b) Where appropriate, use less unexplained network analysis language; e.g. the "nodes" can be "workers" or "persons", and "edges" and "dyads" could be "social interactions" or something like that.
   (c) Explain what the covariates (GWESP, in and out stars, but also team match, metric distance etc.) mean.
   (d) p 10: "Table 5 shows the fitted model where the nodes are added in the order of their average usefulness, as reported by the other nodes". What does this mean? Does it mean that you used some non-uniform choice for p(s) for these estimates?
   (e) Give a clearer explanation of the goodness of fit procedure (in combination with its general explanation, see 2(g) above). What does "the goodness of fit on the in-, out-degree and ESP distributions" (p 11) mean?
   (f) The six goodness of fit plots take up quite a lot of space in the main text. Perhaps consider putting some of them in the appendix.
   (g) p 16-17: The language of "performance in fit" and "simulation" is a slightly confusing way to refer to estimation times. Here it will also help if the methods of estimation for both models have been more explicitly explained before (see 2(a) above).

6. Please summarise briefly in the main text what criteria you used to decide on the questions in Table 8. In particular, discuss this for question 8.

7. For each of the examples in the Appendix, please give also a brief summary (couple of sentences) of the substantive research questions and conclusions, in the language of the application. The current descriptions are a bit lifeless, and give the impression that the purpose of social network analysis is to fit network models rather than to use them to answer substantive questions.

8. Please provide example code for fitting the two kinds of models. This is particularly useful in a comparative paper like this. It can be included as supplementary materials or even (if short and simple) in the appendix. Also, in the "Access to data and computer code" statement on the manuscript submission system (see also the instructions to the authors) provide links to the seven examples for which the data *are* publicly available (as mentioned on p 6).

9. Other comments:
   (a) In the title, I suggest you add "…for social network analysis", or words to that effect.
   (b) "dyad" and "dyad independence" are never really defined
   (c) p 4, l 13: what is s? [element of script-s]
   (d) p 4: what is MOM? PMF?

(e) p 5 l 25: "heavy lifting" is a bit too casual

(f) Include the references to the studies in Table 1. You can then delete Table 10. Also, include the references in the subsection headings of the appendix.

(g) p 8, l 39: This paragraph (which is about relative advantages of LOLOG vs. ERGM rather than construction of your ensemble) does not seem to belong to this place in the text.

(h) p 10: "when substituting the GWESP term for a triangle term which is not possible with ERGM" – I think this says the opposite of what you mean it to say.

(i) p 16: "role", not "roll"

(j) "[!h]" in Table 8, "?? (Add)" on p 21, "?? (Dis)" on p 24, the first two entries in the References – please see comment 1 above.

(k) Table 9: State also in the table (caption or column heading) that 1-8 refer to the numbers in Table 8.

(l) p 19: Delete the lonely subsection number and label for 6.1.