Review of JRSS-SA-Nov-20-0251: Clark, D. A., & Handcock, M. S. "Comparing the Real-World Performance of Exponential-family Random Graph Models and Latent Order Logistic Models"

In the following, citations are to references in the manuscript, with any additional references not in the manuscript at the end of this report.

Summary:

This manuscript compares the well-known exponential-family random graph model (ERGM) with the recently proposed latent order logistic (LOLOG) model (Fellows, 2018b) on a (substantial: 35 end up included in the final analysis) set of networks with ERGM models published in the Social Networks journal. The ERGMs are reproduced (where possible) and then equivalent, and sometimes improved (in terms of goodness-of-fit) LOLOG models, are also estimated.

A short description of the LOLOG model and its relation to ERGM is provided, briefly recapitulating the original descriptions in Fellows (2018b).

The case of the networks from one particular paper (Sailer & McCulloh 2012) is explored in detail, while the rest are summarized in a table with some details in the Appendix.

It is shown that, generally, the LOLOG is qualitatively in agreement with the ERGM. Often the LOLOG is able to fit with simpler terms such as star and triangle, which usually result in near-degeneracy in the ERGM, requiring the "alternating" or "geometrically weighted" model terms instead. Further, the LOLOG tends to be easier (and faster) to fit, requiring less "tuning" than ERGM (where, as mentioned, simple terms like triangles often lead to degeneracy, and decay parameters for geometrically-weighted terms need to be manually tuned, or estimated by curved ERGMs, the latter often being very slow or hard to converge).

The ERGMs are estimated with the well-known statnet ergm R package, and the LOLOG models with the R package lolog provided by the the creator of the model (Fellows, 2018a).

General comments:

The manuscript is easy to read and clearly is the result of significant effort in collecting and re-estimating (in ERGM and LOLOG) models for the ensemble of networks. It makes a significant and novel contribution in taking the "data-centric" perspective (given the data, what modeling approach is best?) rather than the usual perspective of a "methodology" paper (given our model, what data can we fit with it?), showing the usefulness of the LOLOG model for situations in which ERGM would be the conventional choice.

Another interesting contribution this manuscript makes is the suggestion (p.6 lines 30-35) of a reason as to why LOLOG does not suffer a similar degeneracy problem to ERGM. This was not clear (at least to me) from the Fellows (2018b) paper (unlike for example Fellows & Handcock (2017) where phase transitions are explicitly removed by the "tapered" Gibbs measure).

The argument that this population of networks is biased towards ERGM and hence provides stronger evidence for the value of LOLOG models (a conservative indicator that LOLOG is useful) is a good one, given the the "file drawer" effect of only models with good ERGM fits likely to published in Social Networks, and the occasional (or perhaps frequent) difficulty of estimating ERGM models leading researchers to omit terms they might have wanted or considering fewer models. This problem has been discussed also by (for example) Martin (2020) [and his statistics textbook he cites therein].

The only weakness in this argument is if we consider the possibility that LOLOG is a model that works well exactly when ERGM works well (and indeed theoretically this is not so ridiculous, given they are equivalent in the dyad independent case). In which case, this sample is not quite such a conservative indicator of the utility of LOLOG.

That said, we know that this is not the case, given Fellows (2018b) describes LOLOG as "complementary to the popular ERGM framework", and states that there are simple LOLOG models where an ERGM requires prohibitively many terms, and also vice versa. Something like this is also shown in this manuscript, with not all networks being able to be fit better by LOLOG than ERGM, and varying results as to which model class (including nonzero counts for "Both" and "Neither") is "more useful subjectively".

Although I had read the Fellows (2018b) preprint, and was aware of at least one use of the LOLOG model which was useful in a case where an ERGM was not able to be fit, I had not experimented with the LOLOG model myself, until reading this manuscript as a reviewer. This motivated me to try the lolog R package (Fellows 2018a) on some networks which are difficult to fit with ERGM.

The very first example I tried did indeed demonstrate the usefulness of LOLOG: the E. coli transcriptional regulation network (Salgado et al, 2001; Shen-Orr et al, 2002) available in the ergm package (network ecoli1 in data(ecoli)) and for which simple models (inherently directed network converted to undirected; no triangles or other transitivity parameter) are estimated in Saul & Filkov (2007); Hummel et al. (2012). Estimating ERGMs for this network left in its original state as directed, and especially if triangles or alternating or geometrically-weighted versions of transitivity are used, had apparently proved impossible with statnet or PNet/MPNet, and the only such ERGM I am aware of is in Stivala & Lomi (2020), requiring the use of more recently published methods (Byshkin et al., 2018; Stivala, Robins, & Lomi, 2020), which are based on a version of "persistent contrastive divergence" and do not have the provable statistical properties of MoM or MCMLE for example. I was able to easily obtain a LOLOG model of this network including the twoPath and triangles terms (not geometrically-weighted) and in-gwdegree and out-gwedegree (both with decay 0.2 arbitrarily), showing qualitative agreement with the (not exactly the same) models in Stivala & Lomi (2020), with similar goodness-of-fit. Importantly, however, the LOLOG was able to estimate a model with the triangles term, which

ERGM could not (requiring the "alternating" transitive triangles statistic), facilitiating interpretation in this context as a "motif" (although "motifs" are usually considered induced subgraphs, so the relevant test is GoF on the 030T triad, which is good). I was also easily able to estimate a similar LOLOG model for the other network (yeast transcriptional regulation) in that preprint.

Although excluded from consideration from the network ensemble due to the time period chosen (1979 - 2016) and only journal Social Networks, there has been progress in new methods for ERGM estimation for very large networks, and it is now possible to estimate ERGM parameters for a network with over one million nodes, e.g. the ERGM for an online social network estimated in Stivala, Robins, & Lomi (2020). Albeit the ability to estimate ERGM parameters for such large networks relies on the methods discussed above, which achieve this speed and scale precisely by avoiding the expensive simulation to find equilibrium of Markov chain for ERGM, and hence there is as yet no goodness-of-fit test possible for models of such large networks.

By taking advantage of local dependence and latent block structure, hierarchical ERGMs have also been estimated for a network with over 10 000 nodes (Babkin et al., 2020). The latter method is implemented in the R package hergm (Schweinberger & Luna, 2018) and allows the conventional simulation-based goodness-of-fit tests.

These advances in ERGM estimation may somewhat qualify the claim (p. 19, lines 48-50) regarding LOLOG "increasing the maximum size of networks that can be feasibly analysed." Although perhaps LOLOG, with its advantage over ERGM of the ease of simulation from the model, is a better way forward here. It might be interesting to see how well it can scale in practise.

Another aspect of LOLOG, which is used in the online social network example (the one related to hamsters and their owners) in Fellows (2018b), but not used in this manuscript except in one example in the Appendix (due to the selection criteria of the network ensemble) is the edge ordering mechanism. Although it may indeed be "at odds with some researchers' philosophical beliefs" (p. 20 line 31) about networks, the ability to explicitly model edge ordering (and preferential attachment) could be highly relevant (and not so at odds with beliefs about social networks) in modeling citation networks, for example, where ERGMs are sometimes used, and have to include some mechanism to account for citation temporal direction (e.g. McLevey et al., 2018). It seems to me that LOLOGs could be highly useful in such cases, where there definitely is an edge ordering in the data.

Major (i.e. substantive) points [there is only one]:

In the second-last paragraph (lines 42-49) on p. 16, concerning geometrically-weighted degree (GWDEG) term being negative and significant in an ERGM, it says this is "at odds with the LOLOG model". However if I understand correctly, a LOLOG model with a GWDEG was never fit, so this refers to a comparison with a LOLOG model containing the in and out 2 and 3 star terms described in the previous paragraph (lines 28-40). The out-2-star term was found to be positive and significant, the the out-3-star term negative and significant. This is interpreted as a "tendency for super daily interactors". (No similar tendency was found significant for corresponding in-(2,3)-stars).

This finding (positive 2-star and negative 3-star) is a fairly typical result in ERGMs (I am assuming the same applies to LOLOG, although perhaps without, at least to the same extent, near-degeneracy problems as a motivation for geometrically-weighted statistics) indicating a tendency towards centralization ("super daily interactors" here), and indeed the alternation of the signs as higher order terms are added (edge, 2-star, 3-star,...) is a motivation for the alternating-k-[in|out]-stars and gw[io]degree parameters (see e.g. pp. 65-67 of Koskinen & Daraganova, 2013).

However the interpretation of GWDEG (gw[io]degree) in statnet is (confusingly: see Levy (2016); Levy et al. (2016)) that a negative parameter estimate indicates a tendency towards centralization, and positive a tendency against centralization (Hunter, 2007). This is opposite (i.e. signs swapped) to the interpretation of the alt-k-star parameter - and so in fact this would seem to be consistent with the LOLOG out-(2,3)-star interpretation above. Which is reassuring (rather than "concerning"), although the point that the LOLOG allows a better fitting model (or at least more easily to estimate a model that fits well, perhaps not requiring the geometrically-weighted parameters usually required in ERGM to avoid near-degeneracy) stands.

Minor points:

Some typos or missing references, easily fixed e.g.:

"Adolescent Health Survey." and "Pajek Datasets" at top of References (p. 27) and some corresponding '??' missing references (e.g. pp. 21,24), arXiv identifier missing from the Fellows (2018b) reference, etc.

Reference to Table 5.2 (p. 11) should presumably be to Table 7.

"Matching floors play a roll..." (p. 16) should obviously be "play a role..."

One or two LaTeX errors e.g. "[!h]" appearing before Table 8 caption.

Sailer & McCulloch (2012) should be Sailer & McCulloh (2012) [only two "c"s in McCulloh].

References (in addition to those already in the manuscript references):

Babkin, S., Stewart, J., Long, X., & Schweinberger, M. (2020). Large-scale estimation of random graph models with local dependence. Computational Statistics & Data Analysis, 152, 107029.

Byshkin, M., Stivala, A., Mira, A., Robins, G., & Lomi, A. (2018). Fast maximum likelihood estimation via equilibrium expectation for large network data. Scientific reports, 8(1), 1-11.

Hummel, R. M., Hunter, D. R., & Handcock, M. S. (2012). Improving simulation-based algorithms for fitting ERGMs. Journal of Computational and Graphical Statistics, 21(4), 920-939.

Hunter, D. R. (2007). Curved exponential family models for social networks. Social Networks, 29(2), 216-230.

Koskinen, J., & Daraganova, G. (2013). Exponential random graph model fundamentals. Ch. 3 (pp. 49-76) in Lusher D, Koskinen J, Robins G, editors. Exponential random graph models for social networks. Structural Analysis in the Social Sciences. Cambridge University Press: New York.

Martin, J. L. (2020). Comment on "Geodesic Cycle Length Distributions in Delusional and Other Social Networks". Journal of Social Structure 21(1):77-93. doi:10.21307/joss-2020-003

McLevey, J., Graham, A. V., McIlroy-Young, R., Browne, P., & Plaisance, K. S. (2018). Interdisciplinarity and insularity in the diffusion of knowledge: an analysis of disciplinary boundaries between philosophy of science and the sciences. Scientometrics, 117(1), 331-349.

Levy, M. A. (2016). gwdegree: Improving interpretation of geometrically-weighted degree estimates in exponential random graph models. Journal of Open Source Software, 1(3), 36.

Levy, M., Lubell, M., Leifeld, P., & Cranmer, S. (2016). Interpretation of GW-Degree Estimates in ERGMs. Political Networks 2016 (Conference poster). https://doi.org/10.6084/m9.figshare.3465020

Saul, Z. M., & Filkov, V. (2007). Exploring biological network structure using exponential random graph models. Bioinformatics, 23(19), 2604-2611.

Salgado et al (2001), RegulonDB (version 3.2): Transcriptional Regulation and Operon Organization in Escherichia Coli K-12,Nucleic Acids Research, 29(1):72-74.

Schweinberger, M., & Luna, P. (2018). HERGM: Hierarchical exponential-family random graph models. Journal of Statistical Software, 85(1).

Shen-Orr et al (2002), Network Motifs in the Transcriptional Regulation Network of Escerichia Coli, Nature Genetics, 31(1): 64-68.

Stivala, A., & Lomi, A. (2020). Testing biological network motif significance with exponential random graph models. arXiv preprint arXiv:2001.11125.

Stivala, A., Robins, G., & Lomi, A. (2020). Exponential random graph model parameter estimation for very large directed networks. PloS one, 15(1), e0227804.