Peer Community In
Network Science

An application of PCA to social networks and healthy ageing

Steve Lawford *based on peer reviews by* ***Christophe Prieur***, ***Paul Rochet*** *and 1 anonymous reviewer*

Sueur et al. (2024) investigate the influence of an individual's social network structure on various aspects of healthy ageing, including depressive symptoms, life satisfaction, and overall well-being. The primary dataset comprises 73 adults aged 60 and above, residing in the Paris region from 2019 to 2020, who completed a VERITAS socioeconomic/demographic questionnaire; and is augmented with official data on the characteristics of residential neighbourhoods. The authors apply principal component analysis (PCA) to network structure metrics including degree centrality, density, and global clustering, and identify four dimensions that they argue have social significance: homophily, social integration, social support, and perceived accessibility to local services. Unexpectedly, the authors' statistical analysis reveals that none of the PCA dimensions are linked to healthy ageing.

Although network-based PCA dimensions have been used as explanatory variables in other settings, this paper may be the first to apply the technique to healthy ageing. The main result stands in contrast to related literature which indicates that positive social relationships (engagement, sense of community) are related to more favourable mortality and disease outcomes and that these effects persist as people become older. The paper is a useful contribution to an issue that has considerable public health policy importance. It will motivate further research to understand the negative main result, including potential information loss from PCA, issues of small sample bias and identification (relatively few of the respondents were depressed or anxious), specificity of the survey to the Paris region, and more advanced econometric modelling to better understand causal relationships (rather than correlations) between social networks and well-being in older people.

References:

C. Sueur, G. Fancello, A. Naud, Y. Kestens, and B. Chaix (2024) The complexity of social networks in healthy aging: novel metrics and their associations with psychological wellbeing. OSF Preprints, ver. 3 peer-reviewed and recommended by Peer Community in Network Science. https://doi.org/10.31219/osf.io/j9uz8

Reviews

Evaluation round #2

Authors' reply, 29 January 2024

Download author's reply
Download tracked changes file

Decision by Steve Lawford, posted 23 January 2024, validated 27 January 2024

Minor revisions to text

Dear authors,

Thank you for your careful revision. In my opinion, this satisfies the majority of the referees' concerns on the first version of the paper. I have read the revision in detail and have the following minor comments (typos, references, formatting). I look forward to receiving a revision that deals with these.

Best regards,

Steve Lawford

Roucolle et al (2020) use a similar PCA-based network measure approach on air transport data, and this work should be included in your references: https://enac.hal.science/hal-02616818

25/ "negatively linked with the level of study" is unclear
26/ as -> and; was -> were
29/ was -> were
39-40/ "although perhaps less consistent" is not clear
40/ that -> those
41/ "repeated evidence" is unclear
51/ "built environment" is unclear
57/ deleterious -> detrimental
61-124/ break into two paragraphs
76/ comma after "members"
95/ wrong -> incorrect
100-103/ unclear; especially "the new method is compared to another recently introduced approach"
111 and elsewhere/ "graphlets" are also referred to as "network motifs" in the literature (more common)
115/ 25+ -> "25 and above"

120/ "users" is unclear

125/ remove "studies"

140/ 'where' is unclear

151-184/ include standard reference(s) to network measures e.g. papers/books by Newman and others

221/ table -> Table

224/ remove "to"

226/ table -> Table

247/ table -> Table

249/ table -> Table

253/ table -> Table

260 and 261/ include "the" between "and" and "clustering"

263/ corresponds -> correspond

301/ check spelling (accents) on authors' names

397/ At -> To

424/ it's -> it is

431/ dimension -> Dimension

432/ dimension -> Dimension

434/ dimension -> Dimension

453/ as -> and; was -> were

468/ check Agnete... reference (first, last names)

479/ was -> were

536/ it's -> it is

555 onwards/ check all references carefully for one-to-one mapping with main text; and include publisher for all books; e.g. redundant references? lines 575, 617, 720, 739, 776

788/ remove full stop

789/ models -> model

792 onwards/ improve formatting of all tables e.g. columns too narrow in Table 1 (% sign); Solo -> Single in Table 2; Putnam (1995) missing from References; strenght -> strength after line 796 (twice); need to align column headings and values in table after line 796; add full stop at end of line 801

Evaluation round #1

Authors' reply, 09 November 2023

Download author's reply

Decision by Steve Lawford, posted 23 August 2023, validated 24 August 2023

Reply / round 1: Structure and drivers of social networks and their links with health in older adults (by Sueur, Fancello, Naud, Kestens, Chaix)

Dear authors,

Thank you for submitting your paper to PCI Network Science. The reviewers have provided detailed feedback on your manuscript. Two reviewers express positive sentiments, while one reviewer finds the core questions promising but is critical of the paper's execution. I encourage you to address these comments in a thorough

revision. Please resubmit a revised manuscript along with a comprehensive list of changes. Below I summarize the key points from the reviewers' feedback.

Reviewer 1 ("The paper is well-written..."):

1. Suggests checking robustness by verifying results using original variables instead of regression residuals.

2. Recommends conducting rigorous tests for normality.

3. Proposes replacing long tables with figures for better clarity.

Reviewer 2 ("This paper analyzes the link between social network features..."):

1. Recommends that you explore nonlinear or threshold effects.

2. Suggests analyzing subgroups, particularly individuals with high depression levels.

Reviewer 3 ("This paper addresses two important methodological issues..."):

1. Raises significant concerns about the focus of the work and recommends separating the topics into distinct papers.

2. Emphasizes the need for thorough comparison with the existing literature, if the focus is the statistical independence of network measures.

3. Calls for stronger empirical results and clearer interpretation, if the focus is the relationship between healthy ageing and network structure.

Key additional feedback includes:

1. Consider revising the title of the paper to better reflect the main results.

2. Pay meticulous attention to spelling and grammar.

3. Enhance the presentation of all figures.

4. Provide more context and intuition (e.g., interpretation of eigenvalues, PCA dimensions, specifics of the data sample).

5. Improve the paper's structure to highlight the main contributions and to make the results clearer for the reader.

6. Place the paper firmly within the context of the recent literature.

I wish you success with your revision.

Best regards,

Steve Lawford, ENAC (University of Toulouse)

Reviewed by Christophe Prieur, 10 July 2023

This paper addresses two important methodological issues, one about the statistical interdependence of usual (or unusual) network measures, one about the dependence of these measures with healthy aging. Both these issues are relevant in network science, and in aging / health studies. However each of these would benefit from being treated separately and more in depth.

I will discuss these two issues one after the other, but first express some concern about the overall structure of the paper and of the demonstration. Ideas are mixed linearly with very few separations between steps, smaller or bigger arguments, which is especially true in the 5-page discussion section, but also in the setting of the network method, as i will detail below.

1/ On the statistical dependence of network measures.
Very few state of the art is provided on this half-century-long issue. Some recent development in the particular case of personal networks might be a good start: Vacca, 2020 ; Bidart et al, 2018 ; Charbey & Prieur, 2019.

Taking as central this issue in a paper would imply properly defining and discussing the measures, maybe not as meticulously as in Sosa et al, 2020, but not the way it is elusively done here (in a table put as an appendix). This would bring the authors to argue about using the so-called Simmelian brokerage instead of well-established Burt's measures of structural holes. Simmelian brokerage, defined in Latora et al, 2013 (Journal of statistical physics), is far from having brought a large consensus in network analysis: according to Google Scholar, Latora et al, 2013 is reportedly cited 71 times, among which few works in sociology, if any. If one absolutely had to get rid of Ronald Burt (but why, really?), at least one might refer to him anyway, and why

not consider for instance using Vedres & Stark's structural folds instead? Or other measures that have been more thoroughly studied in the field.

Now once the measures are properly defined and discussed, if the main goal of the study were to assess their statistical interdependence, dealing with a much larger sample would be mandatory. Quoting the article: "we need more formal social network analyses (more quantitative and less subjective) [than Cornwell (2009)]". But the present study relies on 72 networks, while Cornwell relies on a sample of 3 *thousand* networks.

Arguably Cornwell's networks are limited in size by design, but in the present study, the maximum size is 19, which is very limited to draw solid conclusions. In comparison, Vacca relies on six datasets between 119 and 385 egos having up to 45 alters, Bidart et al on 287 egos with up to 134 alters, Charbey & Prieur on two datasets of 3k and 10k egos with up to 350 alters.

Moreover, the distributions of the variables on the 72 egos are hard to infer from what is shown in the paper. The correlation table in Figure 1 is difficult if not impossible to read (even zooming-in shows highly pixelated text, with many numbers missing). I could not even read the network sizes ("degree") to check to what extent the sample is limited (the highest bar in the degree barchart is on a very low degree, which suggests networks not much bigger than Cornwell's).

2/ On the ties between network measures and healthy aging.
Here again, taking this issue as central would bring the authors to discuss more in depth the interpretations of their statistical results. Despite the limitation of the sample, both in size (72) and in network sizes (19), there might be relevant insights from other non-netowk variables, or better, from qualitative material (whether the survey contains open questions for instance).
The way they are stated here however, the interpretations are mostly hypothetical, with very few empirical clues.

Some fruitful references on the matter: Daatland & Lowenstein, 2005, relying on a survey of 6k respondents, provides insightful considerations on relationships (while not using network data) ; Wyngaerden et al, 2019, use personal network measures (ego-betweenness, etc) on 380 egos, with whom they have conducted face-to-face interviews, which provides rich empirical material.

To conclude, i find the two ideas of this paper promising, but i don't agree any of the two claims have been sufficiently proved. Once again, the two deserve a paper per se (and keeping them both together blurs the argument), a better account of the state of the art should be taken, and the reduced size of the sample might be turned into a benefit by switching to more empirically-backed interpretative insights.

Reviewed by Paul Rochet, 04 July 2023

The paper is well-written and worthy of publication in my opinion, under minor revisions.

The author use various statistical tools to identify how various social network metrics linked to well-being (or absence thereof) of elderly are correlated. Some of the results of the analysis are statistically significant.
Specific remarks:
- I don't understand the sentence "... collinearity occurs when the correlations were approximately 0.9 or above". since the value 0.9 is purely arbitrary.
- As a non-expert in health related issues or social relationships, I did not understand what the sentence "people living with ego" (line 25) meant until later on in the paper (line 169).

- The approach that consists in replacing the Simmelian brokerage and clustering coefficients by their residuals from a linear regression of the network density seems a little questionable to me. I would suggest to measure instead of the significance of the unaltered variables in addition to the network density in the multivariate linear model.

- line 227, it would be better to use test for normality (ex: shapiro-wilk) and homoscedasticity (ex: Breusch-Pagan) than to "verify" it graphically. Or at least change the term "verify".
- line 269, what does eigenvalue < 1 mean?
- the information of tables 1 and 2 should be displayed graphically instead of in a table in my opinion.

Reviewed by anonymous reviewer 1, 25 July 2023

This paper analyzes the link between social network features and mental health outcomes in a cohort of older adults in Paris. This study contributes to a developing branch of the literature that tries to identify the influence of social behaviors on mental health, which has been recognized as a crucial public health issue over the last years. While the results do not conform with previous studies, since the authors do not find any evidence for a link between social network features (including the number of relationships) and mental health outcomes (i.e., depression, life satisfaction, and well-being), it may provide valuable insights. It is a good example of a study that should be published because it contributes to an important debate, even though it presents null results. Overall, the manuscript is clear, the survey methodology is interesting, and the text is straight to the point. I do not have many comments, but recommend some minor editing.

Major comment:

1.   One point that could be further discussed is that mental health may not related to network features in a linear way. It could be that certain features matter only with some threshold effect (for example, having no friend at all has a negative effect on mental health, but once someone has at least one friend, they are "protected" from loneliness). Because the sample is small, this might have gone undetected. Looking at extreme cases or separating the sample between individuals with high levels of depression and the rest might help identifying whether this is the case or not.

Minor comments:

1.   The title is misleading, as the authors do not examine the "drivers of social networks", and they also do not look at health but specifically at mental health. It would be good to reformulate it to fit more closely the content of the paper.

2.   The manuscript should be proofread. There are quite a few spelling mistakes and grammatical or syntactic errors. Examples: line 13: "Social network is an important factor", line 60 "social networks is the webs", line 72 "network size interacts with personal cognitive and physical decline"(this sentence does not make sense to me), line 127 "our study felt during the Covd-19", etc.

3.   The fourth dimension of the PCA seems difficult to interpret as it is. I would consider either not interpreting at all the results for this dimension or try to understand further what it represents.

4.   Considering the size of the sample, it seems possible to investigate specific cases more closely. In particular, the authors could examine (qualitatively) the cases of the individuals who reported high levels of depression.

5.   The specific context (Paris) may explain the counter-intuitive results of this study, as the authors note line 415. For readers who do not know the city, it could be good to explain this point a bit more.