Comments

This paper has a clear structure and is very well written and could be used as a textbook example for computational analysis of cuneiform texts, or presented as a poster for demonstrating the methodology. The aims and the workflow are described understandably and it's especially nice that the parameters for the Pmizer and Gephi algorithms are explicitly mentioned to make this study reproducible (assuming that the given dataset was distributed with the paper).

Studying relationships between locations mentioned in NA texts is a fresh idea, and well suited for the methods learned in the summer school. This is a stereotypical example of a study that aims to "see forest from the trees", that is, investigating potentially interesting patterns in a large dataset without having to spend hundreds of hours to close-read it and attempting to keep track of thousands of connections. It is yet unclear to me, to what kind of new discoveries linking place names could lead to, but regardless, the presented method is well-suited for discovering and visualizing general distribution of otherwise obscure locations (i.e. is the place X most likely in Urartu, Babylonia or Assyria?), which is very helpful for getting a general idea of the Neo-Assyrian landscape.

The results are communicated and discussed clearly and the author acknowledges potential flaws in the data set. The suggestions for refining the dataset further to disambiguate between KUR 'land' and URU 'city' and to include GIS data to the study are very good ideas for improving the results.

Plotting the network of locations on a map instead of an LNA graph would be a very interesting experiment with a carefully selected dataset, especially if the actual coordinates of some place names were purposefully left out. This would allow us to measure how close to their real known locations we could posit them only on basis on their distribution in the texts (i.e. if our mystery place X has tie strengts of {A, B, C} with places W, Z, Y, it would be interesting to see how the place X posits in relation to these, if we could anchor our reference locations W, Z, Y to their known coordinates and let the gravity algorithm find the best location for our mystery place in relation to these). It's unlikely that we could find the location of Agade using this approach, but it would definitely be an interesting experiment.

Comments

You submitted a very good paper that indicates that you learned a lot in the summer school. Your aims are clearly defined and the paper is well written, logically structured, and shows that you have a critical eye as to the quality of the data. As is also stressed in the paper, this is important and a great strength.

Tools and methods

The earliest stages of the data collection and cleaning processes are only briefly mentioned, but you provide the reader with useful links and literature and as soon as you get to the case-specific methodology, you explain your steps in greater detail; specifying the parameters you used, briefly explaining their relevance and illustrating points with screenshots. Importantly, this makes the process transparent and reproducible, and signals that you made an effort to understand how the code and programs works.

I do not expect you to have noticed this, but as a tip: Gephi warns that the 'prevent overlap' function should not be used with 'approximate repulsion' (if you click on e.g. 'prevent overlap' in the 'ForceAtlas 2' configuration window, you get a brief description of the relevant tool). It is not a major issue for this project, but worth keeping in mind in case you want to do larger LNA or SNA projects in the future.

Your use of GIS to illustrate a discussed point under 'results' is also commendable.

Results and discussion

Your network visualization is informative and you again describe nicely what you did while briefly explaining the main point of the applied tools. It is also good that you include a reference to the modularity metric you used (crediting developers are important and in other contexts, you might have also wanted to include the papers to cite for ForceAtlas2 and Gephi in your bibliography).

The tools you used to highlight specific things in the network graph are appropriate and indicate that you already gained a good understanding of the basics of SNA and of Gephi's functionality. For example, representing the weight of the edges by a black to red scale works well. However, are you sure that '*in practice this shows the total number of connections*'? The 'degree' and 'weighted degree' centrality measures count the connections of the *nodes* (while ignoring or taking the weights of the edges into account), but the *edge weights* themselves usually remain what you input – which I believe are the PMI scores in this case?

Since you conveniently shared your Gephi project via GitHub, I played a bit with the parameters myself. In particular, I found unticking the 'normalize edge weights' useful for seeing more clearly that the model consists of a number of connected components. Since all nodes are positioned relatively close together in Fig. 3, this is not as easily recognized there, so I was pleased to see that you also noticed and explicitly comment on it. For the cohesiveness of the paper, this point works well as a reminder that the dataset has issues and a link to your suggestions for improvements and extensions. For the evaluation of your work, it serves as (one of many) signs that you put a lot of work into this assignment, and that doing so paid off.