Faculty of Sciences and Bio-Engineering Sciences
Arthur Chomé - 0529279

Assignment 3: Review

1   Introduction

Based on our student number (0529279), we had the option between paper 2, 4 and 9 to write our review about. We went for paper number 4(Dandekar, Goel, & Lee, 2013) discussing the issue of opinion polarisation within society and a mathematical model as to measure to which degree opinions of a given set diverge. Thus implying polarisation.

It starts by explaining the factors that lead to this phenomenon which include homophily –interaction between like-minded individuals– and biased assimilation whereupon individuals –when confronted with inconclusive evidence on a complex issue– draw undue support to their own position on the matter, are the leas. After a –quite misleading– introduction on the issue of polarisation and their contributing factors, it refers to DeGroot's model(DeGroot, 1974) for opinion formation criticizing it for its lack of inclusion of biased assimilation. Its contributions are an improved model with biased assimilation applied on a network model of individual opinions –linked to each other if the persons know each other– taking the form of a directed graph. It also analyses 3 recommender algorithms that present data over the internet to a user based on his own convictions and how this may possibly lead to polarisation.The paper gave us a good review of the mathematical possibilities to measure polarisation within a network of opinions linked to a larger group. It's a surprise to read that even the most complex social problems can be mathematically expressed as models and explained using graph theory.

2   Opinion

The paper's subject is good –trying to mathematically measure the social problem of polarisation– but it lacks the necessary touch to make it attractive for political or social scientists who would profit the most from this article. Its introduction is bad discussing polarization and all additional concepts but the content following upon it goes into advanced mathematical domains without teasing it beforehand making it very misleading and confusing.

If we were only to base ourselves on the introduction of the paper –its writing style and title– we might convince ourselves that it was written for a target audience with no explicit mathematical background such as political scientists or psychologists. However, the paper shifts directly towards a mathematical focus introducing rather complex notions about graph theory and models making for a bad writing style. Another fault of the paper is has a bad overall structure. Though it provides an introduction, the separation between related work, experiments and conclusion is not sufficient.

The model uses three recommender algorithms as to measure to which extent they cause polarisation on the internet when biased assimilation is present: SALSA(Lempel & Moran, 2001), Personalized PageRank(Page, Brin, Motwani, & Winograd, 1999), and item-based collaborative filtering(Linden, Smith, & York, 2003) respectively. Though they are properly referred, they only provide mathematical formulas to explain the workings of each algorithm making it very

hard to follow. Another problem is that the link between DeGroot's model and the analysis with these recommender algorithms lacks clarity. A summary on the tests being performed that indicate increased or decreased polarization is missing making it harder to believe these claims. It mentions a graph of opinions to apply the model on but does not elaborate on this. It does not show any example data which makes it hard for the user to follow.

3   Improvements

For the above critique on the paper, we have some suggestions for improvements to be made. This would in turn increase the relative attractiveness of the paper increasing its influence. The following suggestions are focusing on the writing the style, the paper's structure and the discussion of algorithms.

3.1   Content, Writing Style & Structure

The paper's introduction focuses too much on the terminology and concepts of polarization. It is misleading to be afterwards discussing the paper's extension of DeGroot's mathematical consensus model for the explanation of divergence in opinions. A suggestion is to shorten the paper's introduction on polarization and focus more on the mathematical model they developed for it. The title of the paper should be reworked, it does not stick with the actual content of the paper. After the introduction, the paper is full of definitions and formulas on mathematical models making it very difficult to follow. I suggest to remove the definitions in the paper in exchange for a more high-level, easy-to-understand explanation all the while referring to papers citing these definitions. The use of figures should not be excluded.

Finally, the overall structure of the paper should be reworked with a clear related work section about DeGroot's model, a discussion about the used node graphs and the Network Disagreement Index. After this comes a discussion of the extension of the model and the experiment with recommender algorithm followed by a clear conclusion.

3.2   Overview of Analysis

Providing an Appendix on the results of the polarizing nature of recommender algorithms is a good start but it should be integrated within the paper as to be more clear for the readers. It should give a hands-on comparison of the scores for each algorithm and clarify which difference it would yield if we were to use DeGroot's algorithm without biased assimilation.

As to verify the result of the model, the experiment could be recreated in real-life with a sample population. Recommender algorithms work to provide customized news feeds to each user and the internet does not lack API's to recreate these situations in a measured environment.

3.3 Algorithm Performance Comparison

The inclusion of pseudocode[1] for each algorithm should not be missed. It provides a high-level description of each of their operation. Although all three algorithms operate by ways of random walks on a node graph G, there is a performance comparison to be made. One of the biggest concerns in Computer Science is to perform work as efficient as possible without wasting resources. This is why a table of the best, worst and average case performance of each algorithm should be provided in Big-O notation[2]

References

Dandekar, P., Goel, A., & Lee, D. T. (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, *110*(15), 5791–5796.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), 118–121.

Lempel, R., & Moran, S. (2001). Salsa: the stochastic approach for link-structure analysis. *ACM Transactions on Information Systems (TOIS)*, *19*(2), 131–160.

Linden, G., Smith, B., & York, J. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*(1), 76–80.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The pagerank citation ranking: Bringing order to the web.* (Tech. Rep.). Stanford InfoLab.

---

[1]`https://en.wikipedia.org/wiki/Pseudocode`
[2]`https://en.wikipedia.org/wiki/Big_O_notation`