

Data Mining for Online Social Networks – Review

I have annotated the document where I have spotted typos or formalities that are wrong. Apart from that, I still have some general remarks:

When defining the neighborhood of a node you should be careful to mention that those are only the vertices that can *directly* be reached from that node, not any vertex that can be reached.

Perhaps you should briefly mention what lattice graphs are. Also, you state that "random networks often have this property, whereas lattice graphs usually have a high CC". I think that this formulation is a bit unclear and you should explicitly state that lattice graphs have a high CC, but a low L .

I suggest you also explain why the small world network property is important. Additionally, the construction of networks with that property feels a little bit out of place since—from my understanding—we usually are given a graph and don't want to construct one.

You describe that snowball sampling suffers from "boundary bias", but don't explain what that is.

What exactly is the step response in an LTI system?

You introduce the concept of modularity with a matrix E , where " e_{ij} represents one community with the value being calculated by the fraction of all edges connecting community i with community j ". What fraction are you referring to exactly? The only number I see is the amount of edges connecting the two communities, so what is this divided by?

As I have already described in an annotation, it is important that you explicitly refer to every figure in the text. Also, the figures should only support your argument and never be needed to understand the paper. Thus, it might be a good idea to describe the essentials of Fig. 10 in the text, before referring to it.

All in all however, I like the structure and content of your paper. There is still some room for improvement as outlined above and there are several minor typos and formalities that you should fix, but otherwise I think that your paper gives a very good introduction to data mining in OSNs.