Social and economic networks. Final exam. 2nd part

(It is enough to solve exercises for an amount of 10 points)

Read the paper "Social triangles and generalized clustering coefficient for weighted networks" by R. Cerqueti, G. Ferraro and A. Iovanella (arXiv preprint arXiv:1712.01561 (5 Dec. 2017), `https://arxiv.org/abs/1712.01561`), and then answer the following questions.

1) In a single sentence, what is the main question addressed in this paper? Why do the authors consider this question relevant and worth answering? After reading the paper, do you consider their answer relevant? Why?

2) How does this paper relate to previous work about the question addressed in it?

3) What are the main definitions in this paper?

4) On your behalf, what are the strongest contribution and the greatest drawback of this paper (if any)?

5) Do you consider the numerical experiments provided in Section 5 suitable? (I'm not referring to the networks analyzed, but to the computations performed on them.) What conclusions do you extract from these experiments? Do they agree with the authors' conclusions? What other experiments (on the same networks) would you perform?

6) In their experiments, the authors mostly consider cases when $\alpha = 0$ or $\beta = 0$. How would you choose α and β so that only really meaningful pseudo-triangles were considered?

7) The authors define their generalized clustering coefficient only for undirected networks. How would you generalize it in a sound way to directed networks? Provide a definition that makes sense on some class of weighted directed networks of your choice with some specific meaning of its links (for instance, non-symmetric friendship networks, trade networks, citation networks with weights the number of times an author has cited another author, ...)

8) At the end of the day, the final value of the generalized clustering coefficient does not depend on the specific weights of the edges in the triangles and "generalized triangles": each such a triangle only counts 0 or 1 in the definition. Can you think of some sound way to use the weights to "weight" the triangles and "generalized triangles" in the definition? How could your definition change the experimental results explained in the paper?

9) How is the generalized clustering coefficient related to the Strong Triadic Closure Property? Can you find a way to test whether a network satisfies the STCP using this coefficient (together with other indices, of course)?

Now, some technical questions:

10) In page 4 the authors say "Our generalized clustering coefficient has a further very relevant property: it assumes unitary value in several situations and not only when the graph is a clique." Is it true that if an unweighted graph has clustering coefficient 1, then it is a clique? And if we assume moreover that the graph is connected? Justify your answer.

11) In equation (3), page 7, the authors recall Barrat et al's clustering coefficient for weighted networks. I also introduced this index in my lectures, but equation (3) and my definition do not define the same number. What is the difference? Which one corresponds to Barrat et al's original definition? How did you check it? If the correct version is the one given in this paper, I probably had a good reason to change it. What was this reason? And, if the correct version is the one I gave in the lectures, did the authors of this paper have a good reason to change it (for instance, to compare it with their proposed coefficient)?

12) In equations (4)–(5), page 8, the authors recall Zhang—Horvath's weighted clustering coefficient for gene co-expression networks. Does it fit with the general philosophy of Kalma–Highman's indices or Opshal–Panzarasa's indices as explained in my lectures?

13) (Extra bonus) The authors mention that they used R to implement their coefficient and perform their experiments, but they do not provide their implementation. Implement in R their generalized clustering coefficient as well as your proposals in points (7) and (8), and repeat their experiments with your functions, or perform new experiments (for instance, using other networks, using your bounds in point (6), checking point (9), ...).