Answers to the Report for TCSS-2017-12-0106

Fatima K. Abu Salem[1], Mohamad Jaber[1], Chadi Abdallah[2], Omar Mehio[1], and Sara Najem[2]

[1] Computer Science Department, American University of Beirut, Beirut, Lebanon.
Email: {fa07, mj54, okm02}@aub.edu.lb
[2] National Center for Remote Sensing, National Council for Scientific Research (CNRS), Riad al Soloh, 1107 2260, Beirut, Lebanon.
Email: {snajem, chadi}@cnrs.edu.lb

Thank you very much for your letter regarding manuscript TCSS-2017-12-0106. Following the insights provided by the reviewers, we have taken into account all remarks and answered the questions. In this letter, we provide the answers (in italics) to the comments of the reviewer. We also attach herewith the modified version of the paper. We have highlighted in red the modifications directly affecting the manuscript. Additionally, we highlighted in red the section titles of completely modified sections.

We would like to take this opportunity to express our sincere gratitude to the reviewers and editors for helping us significantly improve on the quality of our paper.

Sincerely Yours,

F. Abu Salem, M. Jaber, C. Abdallah, O. Mehio, and S. Najem

1 Answers to the comments of Reviewer 1

I do think the study of vulnerability analysis of the power grid network is very attractive and very useful. However, there are two major things I felt this paper is not quite enough for publication quality.

- The network is too small and it does not make sense to use distributed systems.
- The method in this paper is specific to the studied network that is not general enough. Although the authors provide arguments for both in the response, but the arguments are not weak and not convincing. Since the network in study is pretty small, I would suggest the authors to improve their paper from the two aspects: (1) focusing more on the vulnerability analysis; (2) in parallel/distributed aspects, since the graph is small, it would be better to parallelize the betweenness centrality for a connected component since that's the main computation, although there have already been plenty of studies on parallelizing the betweenness centrality algorithms in the parallel computing research community.

- *Although the initial input graph is not large, the vulnerability analysis requires (1) very heavy computations and (2) to build intermediate graphs, which is the main reason to use a parallel distributed framework such as Apache Spark. Moreover, benchmarks show that a distributed implementation is required as it is scalable with respect to the size of the graph.*
- *We elaborate more on the vulnerability analysis especially spatial analysis and its importance in our context;*
- *As (1) the graph structure consists of numerous connected components where most of them are of small sizes; and (2) the connectivity can be computed and aggregated concurrently between components (Propositions 1-4), in particular, updating the betweenness centrality scores after reach removal in the cascading scenario requires information about the paths connecting each vertex to vertices only in its own local component and thus can be performed independently of other threads; then using a dedicated parallel implementation of betweenness centrality would not be very beneficial. Add to that, as we are using RDD part of the implementation to compute local betweenness centrality is automatically parrallelized. On another hand, for large components, we can easily use and integrate existing specific parallel implementation of betweenness centrality*

2 Answers to the comments of Reviewer 2

Recommendation: Accept as Regular Paper. No further comments.

Thank you for your appreciation of our work.

3 Answers to the comments of Reviewer 3

Recommendation: Accept as Regular Paper. Summary of Evaluation: Excellent.

- The authors assume that the graph is undirected. Is it the case of lebanese grid system? What is the impact on the performance if the graph is directed?
- Section D. You mentioned that: The spatial correlation analysis shown in Fig. 20 reveals a, ..., which is in agreement with the literature results in [19]. Could you please elaborate more that? Is there a correlation between spatial analysis and loss?
- When computing the loss, you assume that all the nodes and connections are identical. Practically, is it the case?
- The authors briefly discuss the spatial correlation analysis results, concluding with a very brief statement that the failures are spatially long-range correlated. Whilst this is a mathematical conclusion made clear by their computations, I fail to see the real implications of such a result. The authors need to elaborate on that part.

- *Yes indeed. The Lebanese grid system as obtained by the ministry of Power is undirected. Please note that this bears no impact on the performance if the graph is directed. First off, all of the algorithms used (e.g. SCC and BC) have variants that can tackle directed graphs. Moreover, these variants have the same work complexity as those for the undirected graph.*
- *Not really. There is no such correlation. However, the reason we undertake this spatial analysis is because of the following. The propagation behavior of cascading failures remains unknown, unless one performs the kind of spatial analysis that explores the data for such behviour.*
- *Not really. However, the fact that SCC and BC address structural attributes of the graph rather than the weights associated with the connections or the actual nodes, make this assumption a safe one. This assumption is also valid in some of the seminal the literature cited by our manuscript. Please refer to references 4 and 19.*
- *Overload failures usually propagate through collective interactions among system components. It is important to understand the spatial correlation between distance and failure. Our results reveal that high failures in critical nodes has impact that propagates across long path lengths on Lebanese soil.*