

Response to Committee Comments on Dissertation

Sarah LaRocca

June 30, 2014

1 Comments from Ben Hobbs

- **I don't like the endnote approach to references (author date easier to read and refer to). But you can stick with it if you prefer (however, need to distinguish footnotes, see below).**

I have left the reference style as-is because I used the bibliography style (IEEE) included in the JHU latex template for dissertations.

- **The paragraph on p. 4 should be divided into 3 (too long).**

This paragraph has been divided.

- **p. 9 – When a table or figure is called out in the text, it should immediately follow (that page or the next, not two pages later). This is particularly annoying in Ch. 2.**

This has been corrected.

- **p. 10. 25 presents a summary of degree distribution parameters. Should be Authors name²⁵. (Like Crucitti et al. on p. 8)**

This has been corrected.

- **Table. 1.3. What is the asterisk before D for (*D)?**

The asterisk has no meaning and has been removed.

- **Should 1.3.3 be called summary of network robustness modeling, given your earlier differentiation?**

This change has been made.

- **1.4.1 “electric power systems it of utmost” (sic)**

This has been corrected.

- **voltampere should be volt-ampere throughout**

This has been corrected.

- **30. stray references in middle of page**

This has been corrected.

- **43. space missing between (2010) and et al. for Bobbio**

This has been corrected.

- **47. “robustness if limited.” (sic)**

This has been corrected.

- **Replace “(p < 2.2e - 16)” type statements in text with 2.2x10-16 (with -16 as exponent)**

This has been corrected.

- **Chapter 2: please report in a table for all regressions the values of all coefficients (intercepts, slopes), their units, the standard errors for all units, and the R^2 (and adjusted R^2). That is you should report enough that others can easily apply your equations to their networks, and others should be able to assess their significance. Also helpful in assessing the importance of different variables is a report of the mean value of each independent and dependent variable. The figures by themselves are not enough for a thesis.**

The mean values of each independent variable are already reported in Figures 2.1 to 2.4 and in Table 2.3. The mean values of the dependent variable are already reported in Figures 2.5, 2.10, 2.14, 2.17, and 2.20, and have been added in Table 2.4. Regression parameters have been added in Tables 2.5-2.9. Pseudo R-squared values are presented in Table 2.10.

- **Equation (2.10) what are values for targeted attacks? (You mention that you’d expect them to be higher)**

These values have been added to Section 2.3.2.1.

- **p. 89. Right quote mark missing in footnote**

This has been corrected.

- **In addition to Overbye paper, worth citing Baldick’s paper and Stott’s paper: Stott, DC Power Flow Revisited, Power Systems, IEEE Transactions on (Volume:24 , Issue: 3) 2009; Empirical analysis of the variation of distribution factors with loading, R Baldick, K Dixit, TJ Oberbye - Power Engineering Society, 2005**

The Stott paper is already cited. A reference to the Baldick paper has been added.

- **In 3.2.4, worth mentioning that dynamic models (of stability) are a subclass (as opposed to analyzing cascading outages by a**

sequence of steady state models)

A paragraph mentioning this has been added to Section 3.2.4.

- **Define logit link function, and any other functions in beta regression**

This definition has been added.

- **96. “i.e.,r” missing space**

This has been corrected.

- **3.3.2.3 When et al. is at the end of a sentence, you need two periods “et al..”**

According to the Chicago Manual of Style (<http://www.chicagomanualofstyle.org/qanda/data/faq/topics/Punctuation/faq0040.html>) and several other sources, two periods should never be used in a row, even if the sentence ends in an abbreviation. Therefore this change has not been made.

- **105. Indent missing in last paragraph**

This has been corrected.

- **106. Endnotes and footnotes should always follow period (line 3). also section 3.3.4, and elsewhere in this chapter. Because footnote and endnote have the same format and can be confused, I suggest superscripting the full text “Footnote 3” (rather than “3”) so people don’t think it’s a reference.**

The placement of footnote superscripts has been corrected. In addition, citations are no longer superscripted, in order to differentiate between citations and footnotes.

- **Figure 3.2. It’s helpful to have an error statistic as well (since the points are so dense), such as a RMS shown on the graph. The cloud might hide, for instance, that 1-deltaENE does better than 1-ENE (depends on the density of points, which is hidden) Font on Fig. 3.5 can’t be read, please enlarge. Whiskers are pretty useless because of the overlap. Figure 4.1– same problem**

Figures 3.4 and 3.5 have been edited for legibility and whiskers have been removed. Error information is now included in Tables 3.5 and 3.6. Figure 4.1 has been enlarged.

- **Stray footnote at start of p. 120.**

This has been corrected.

- **Footnote 10’s format for quotes (single) is different than for similar footnotes in other chapters.**

They were actually all single quotes but have been changed to double quotes.

- **Footnote 6: period missing at end of sentence**

This has been corrected.

2 Comments from Jonas Johansson

- **Page 3, Paragraph 2, Line 6 Add “at system level” to more clearly differentiate between component/system reliability**

This change has been made.

- **Page 6, Definition 4**

Is your meaning that resilience is only considered by the recover perspective, i.e. not the initial performance degradation after disturbance? How you use it in the thesis (only at two places I think) I believe that the latter perspective is also included. Consider changing, but if not changing it make sure that you use the term as defined.

The two references to resilience have been removed given that their use did not quite fit with my earlier definition and did not add anything to the text.

- **Section 1.3, Page 7, last sentence**

Clarify in what respect physical data of networks is not available, e.g. not available for researcher or lacking in general? All you need is to add a short description, e.g. adding a parenthesis with brief description.

This clarification has been added.

- **Page 9, Amaral reference**

Since you are directly referring to Amaral, it might be better to use the three classifications he provides: (a) scale-free networks, (b) broad-scale networks, (c) single-scale networks. In parentheses you could then add (power-law), (power-law with cutoff), (exponential or Gaussian)

This change has been made.

- **Page 12, Section 1.3.1.3, Heading**

I think it is better to describe it as “Average path length” as you use the reserved “path length” (dij) to calculate average path length. Go through the thesis and change where appropriate (e.g. page 51).

I don't think it makes sense to rename this section, because I don't only use average path length in this thesis (I use max [diameter] and standard

deviation as well). However, I did try to clarify my notation by defining average path length as ℓ_{mean} rather than just ℓ .

- **Page 33, Section 2, Line 2**
“Their modeling approach is then similar to that described above for the OPA model, except they use recalculated values of load to check for exceedances, rather than a power flow model.” Consider adding that Motter and Lai is looking at node overload and OPA (as far as I has understood it) only looks at edge overload. I think this is an important difference that should be highlighted.

This change has been made.

- **Page 38, eq. 1.32**
I am missing an “n” in the equations (as you refer to n in the sentence that follows the eq.)

This has been corrected.

- **Page 39, Paragraph 1, Line 3 after eq. 1.33**
You use the word “risk” to describe ci inoperability. It could be that this is how Haimes and others use it, but does it fit into your definition?

Haimes defines risk of inoperability as “a generic term that measures the joint effect of the probability (likelihood) and degree (percentage) of the inoperability of a system. Risk of inoperability can also be viewed as an extension of the notion of unreliability.” This definition is consistent with my definition of risk.

- **Page 40, Section 1.5.2. Sentence 3**
“However, there is a growing body of research that network theory to examine multiple, interdependent infrastructure systems.” Feels like something is missing, please rephrase.

This has been corrected (the word “uses” was missing).

- **Page 41, Section 2, Sentence 1**
Consider adding “removed” as to “...critical fraction of removed nodes required..”

This change has been made.

- **Page 47, Section 1, Last sentence**
Typo: “if” should be “is”

This change has been made.

- **Page 51, Last sentence**
Consider adding “typically” (or something similar) as to get “These node failure events typically result....”

“Typically” is not needed here because the sentence specifies that the node failure events result in disconnection of *one or more nodes* from the remainder of the network. The nodes removed in a failure scenario are always disconnected from the remainder of the network, so this statement is correct.

- **Figure 2.2, Top-right figure**
Fix the numbers on the x-axis, hard to read.

This change has been made.

- **Page 56, last sentences**
You use 4 criterion to define node importance. Is it ever possible that after this there is still a “draw”. If so, how do you resolve it. Please add a statement of this.

The following statement has been added for clarification:

If two or more nodes are ranked equally, their failure order is selected uniform randomly.

- **Page 64, Sentence 1**
“Higher average nodal degree would imply a network in which connections are concentrated in a smaller number of nodes.” Is this valid for all type of networks? (I would guess it is only valid for the type of network you have generated (power-law with cut-off) – as a network could consist of nodes were all are equally connected to four other nodes, i.e. giving a high mean node degree but does not reflect that smaller number of nodes with many connections.) Please clarify in the thesis.

It is correct that this statement is not necessarily true - it depends on the shape of the distribution of the degree values. This paragraph has been rewritten to remove this statement.

- **Page 66, Section 2, Line 12**
Consider changing “failed” to “been removed” in sentence “The color of a node reflects its degree; nodes shown in gray have either failed or been disconnected...”

This change has been made.

- **Page 67, Section 2, Sentence 2**
“To do this, I remove each node one at a time, and then look at the value of S10 for the modified network based on simulations and separately based on the regression model for 10% of nodes removed.” I think this is a bit unclear (I think I get it though after some thinking), could you rephrase it somehow to make it clearer.

I have rewritten this section to read:

To do this, I remove a given node from the network, and determine the value of S_{10} for the modified network simulating random failures in 10% of the nodes. I repeat this process, removing each node in the original network removed (one at a time). For each of these modified networks, I also use my regression model to predict S_{10} . I then rank-order the nodes by $1 - S_{10}$ for both the simulation results and the regression predictions.

- **Figures in chapter 2**

It would be good if these came in conjunction with the text and not as now at the end of the chapter to increase readability.

This has been corrected.

- **Figure 2.17 and 2.18**

Hard to read the labels of the figure (increase font size)

This has been corrected.

- **Figure 3.4 and 3.5**

Increase font size and maybe log y-axis(?).

This change has been made. The log of the y-axis has not been taken because a few values are very near to zero, so doing so makes the plots look worse.