Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Sajjad Asefi

PhD Program: Engineering Systems

Title of Thesis: Advancements in power system state estimation: innovative algorithms and solutions for enhanced reliability and efficiency

Supervisor: Assistant Professor Elena Gryazina

The thesis document includes the following changes in answer to the external review process.

I am grateful to the jury members for their positive feedback and useful comments. I am happy to address the comments and questions in this document and in the revised version of the thesis.

Prof. Henni Ouerdane

1. The Introduction chapter needs more work. First, there is not enough background information that allows the reader to quickly understand how the research work fits into the broader context of modern power grids. While the objectives have been stated, a proper research question has not been formulated. The main hypotheses that underpin the approach adopted and the methodology, are not clearly stated. The thesis should show how the hypotheses have been tested and the outcomes. Further, although there is a literature review chapter, it would make sense that works are cited and possibly briefly discussed to help with the formulation of the research questions and the hypotheses.
   I highly appreciate this comment. I completely agree that the clear statement of hypotheses beyond the statement of the thesis objectives is missing. In order to address this comment a brief background about the problems have been stated. Then the research gap is pointed out and right after that, the hypothesis has been stated. This will provide an overall understanding to the reader that what is expected to be provided in the thesis.

Prof. Ramesh Bansal

1. Elaborate chapter 2 by adding more recent references from the literature. Include a comparative table showing key points and challenges from the state of the art.
   I highly appreciate your comment. Some new references have been added to the research work. At the summary section of the second chapter (i.e., literature review) a table which summarizes the references has been provided stating the challenges with respect to the discussed topics.

2. Include a summary at the end of each chapter.
   Thanks for your comment. A summary has been added at the end of each chapter.

3. Describe all terms in the equations in the text.

Thanks for your comment. The equations have been thoroughly revised so as to explain all the terms.

4. How does the proposed method prove viable for the case of single bad data in larger networks? Discuss the analysis of both the weighted and normalized residuals in the presence of gross errors when compared to $\chi 2$ testing.
   I highly appreciate your comment. As it has been stated in the thesis, it is assumed that single bad data, or generally bad data can be detected using $\chi 2$ testing. As an example, if we have a gross error in one of the measurement units, it can easily be detected taking into account that the errors follow a Gaussian distribution.

5. How can grid operating procedures be taken into consideration to achieve better data detection?
   Thank you for your interesting question. In order to achieve an improved bad data detection scheme, taking into account the grid operating procedure can be effective. As an example by understanding the typical operating ranges of grid parameters, data detection algorithms can identify data points that deviate significantly from these ranges as anomalies. In our thesis application of forecasting aided state estimation (FASE) to some extend aligns with this concept. Because in case of FASE, based on historical data, system state can be predicted and later utilized as an index to detect anomalies.

6. Considering the expansion of the future grid, what are the other types of anomalies that can be considered as future directions for the development of ADCIT? Discuss in brief.
   I highly appreciate your comment. Other than false data injection attack and sudden load change, taking into account renewable energy sources and their intermittent nature, sudden generation change can be considered as another type of anomaly for future research.

7. Hyperparameters are tuned using sequential optimization in the ADCIT tool, how does it benefit from other optimization techniques?
   Thank you for your comment. Although, this was not the main focus of our research work, but Sequential Optimization is a common and robust method for hyperparameter tuning in machine learning algorithms. It involves systematically exploring different hyperparameter values, evaluating the performance of the algorithm for each configuration, and selecting the configuration that results in the best performance. Simplicity, Convergence, and scalability are some of the benefits of this method.

8. Chapter 6: Explain the feasibility of conducting an economic analysis of blockchain implementation.
   I highly appreciate your comment. In our research work we have provided an initial point for other researchers to implement distributed state estimation easily, but in reality a private blockchain should be used, and using this scheme might require new equipment installation which economically might not be in favor of the utilities.

9. Check the line indent on all pages.
   Thanks for your comment. The thesis has been thoroughly checked for line indent.

10. Check the grammar and tense of the text throughout the thesis. Check line spacing.
    Thanks for your comment. A general overview of the thesis is done to check the grammar

Prof. Qiuwei Wu

1. The anomaly detection is investigated. For the developed method, is it possible to also consider the cyber security issue? is it possible to detect cyber attacks? are there any counter measures to deal with the cyber attacks?

   I appreciate the reviewer for this question. Cyber security is a wider topic and mainly is discussed in the area of information technology. However due to utilization of the communication medium, the power system might also be prone to cyber-attacks. There can be various types of cyber-attack in the field of power system, but one the conventional ones which targets to bypass the bad data detector (BDD) within the state estimator unit in the power system is False data injection attack (FDIA). In this research we have made an effort to detect FDIA when it bypasses the BDD. In the second phase, we have proposed a data-driven method to distinguish it from system typical events, (in our case sudden load change), which can have similar behavior as FDIA.

   As it is discussed in [1], a trivial counter measure would be to identify (i.e., to specify the states or buses which have been target to FDIA). In case of single-state FDIA (means only one state has been targeted), or single-bus SLC (means only load at one bus has been changed), the method is capable of identifying the target buses or states with a high accuracy. While in the case of multi-bus SLC and multi-state FDIA it is still challenging. Due to this fact, the mitigation stage is not at the scope of this thesis and has been proposed as future research direction.

   [1] **Asefi, S.**, Mitrovic, M., Ćetenović, D., Levi, V., Gryazina, E., & Terzija, V. (2023). Anomaly detection and classification in power system state estimation: Combining model-based and data-driven methods. *Sustainable Energy, Grids and Networks*, 101116.

2. The privacy is a very important issue. The blockchain technology is used. How about other methods?

   Thank you for the interesting question. In this thesis blockchain has been utilized to increase the security of the information transfer. There are other methods to enhance the security of data transfer, such as encryption or digital signature, which have blockchain utilizes a combination of them. That's why blockchain is on top of the list, decentralization, data integrity and transparency.

3. Is it possible to test the methods with real measurement data?

   Thanks for asking this interesting question. Although we have not tested our method in real-world yet and the state estimation has been implemented in Matlab environment. But in order to make it more close to the real-world we have inserted random errors in the measurements. These errors follow a Gaussian distribution. In reality the errors can be from different sources, with different probability distribution. However, central limit theorem states that the sample mean of a large number of independent and identically distributed random variables will approximately follow a normal distribution, regardless of the underlying distribution of the individual random variables.

Prof. Oleg Khamisov

1. Section 3.1 is dedicated to power system modelling. Transformers are not considered in this section (which is a logical assumption for this work). However, figures 3-2 and 5-3 have transformers in them.
   Thank you for your comment. In really the IEEE 14 bus system has transformers located on the lines which has been demonstrated in figure 3-2, and figure 5-3. Moreover, these figures show the placement of generation units in the buses. However, in rest of the figures which are devoted to distributed state estimation, to show different areas, the emphasize is mainly on the connection of the lines, and generation and transmission units are neglected within the figures.

2. Formula (3.2). No explanation is given for the entities x_t, z, h and R. What dimensions do they have? Is matrix R positive/negative definite? What does index t mean? If h is a mapping, is it continuous, differentiable etc.?
   I highly appreciate for pointing out the issue. The parameters have been briefly explained after the equation.

3. Text would benefit from the reference for the formula (4.1).
   I highly appreciate your comment. The reference has been added to the equation.

4. In formula (4.6) and other optimization problems what are optimization variables? What sets do these variables belong to (i.e. can x be positive or complex etc.)
   I highly appreciate your comment. "*x*" is the vector of state variables. In the case of AC state estimation is the voltage phasor, which is voltage magnitude and phase angle as shown in (4.12). However, in the case of DC state estimation, since the assumption is that voltage magnitudes are "1", the phase angles are the only variable. For more clarification, the text before (4.36) is slightly modified and after (4.36) the variables are stated clearly.

5. In formulas (4.10) and (4.11) should it be "×" instead of "."?
6. In some formulas (i.e. (4.7)) product is denoted by "×" and in some (i.e. (4.36)) no symbol is used. It is a bit confusing, since "×" symbol is often used for element-wise product. I suggest using one type of notations throughout the thesis.
   I highly appreciate your comment and pointing out these issues. I went through the equations and modified the notations to avoid confusion.

7. Equations (4.13), (4.14) do not have line shunt elements, but equations (4.15), (4.16) do. Please check, model is correct.
   Thank you for your comment. The equations are according to [1], and they are correct, the line shunt elements are considered in calculation of the Y_{bus} matrix (complex bus admittance matrix).

   [1] A. Abur and A. G. Exposito, Power system state estimation: theory and implementation. CRC press, 2004.

8. In (4.10) H is a function of x. In (4.36) after transition to DC equations it does not have any arguments. Please add a brief explanation.
   I highly appreciate your comment and pointing out this issue. For DC state estimation the relation between measurement and the state variables is linear which means the *H* matrix is not related to the state variables, but only to the network parameters. This explanation has been added to the explanation before (4.36) for the sake of clarity.

9. In (4.41) $\forall\, k$ should be $\forall\, k \in K$.
   Thank you for your comment. The equation has been modified.

10. In (5.21) elements can elements $c_{i,j}$ be equal only to 0 or 1, or can there be weights associated with them?
    I highly appreciate your comment. $\boldsymbol{C_L}$ is defined only to represent the connection between the buses in the power system, so its elements has been set to be "1" in case there is a connect and "0" in case there is no connection. The weights of the are assigned using another matrix

11. In (5.22) what W_L is equal to if $c_{i,j} \neq 1$ and i ≠ j?
    I really appreciate for pointing out this issue. The equation has been modified.

Prof. Dmitrii Shatov

1. The section devoted to the blockchain use falls a little out of the general outline of the research aimed at developing methods for state estimation of power system, although in general it corresponds to the main work topic.
   Thanks for the comment. The reviewer correctly observes that the section on blockchain use is to some extent detached from the main focus of the thesis, which is developing state estimation methods for power systems. While blockchain is a relevant technology for improving the security and data integrity of state estimation systems. Due to that reason we have decided to

2. According to Fig. 5-2, in case of bad data detected in the system, the proposed anomaly recognition algorithm is not applied. From this point of view, it is profitably for the malicious attacker to disguise his actions under analogous situation. This case is not considered in this work.
   Thanks for your interesting comment. To my opinion the analogy with such a behavior would be changing simply the observed measurement value from a meter, without coordinating it. So, in this case again the conventional bad data detector would be able to find out the error. Other than that, if the attach bypasses the bad data detector, the proposed methodology would be involved.

3. The classification algorithm contains a tunable parameter γ, which selection on practice is not described in sufficient details.
   I highly appreciate your comment. The details regarding the selection of the hyperparameter is added. More details of implementation of these methods in the simulation environment is provided in chapter 6.

4. There are typos, negligent and illogical moments in the work:

   - the work uses a huge number of acronyms and abbreviations, but the work lacks a separate list of them with explanations to make it easier to read.
     Thank you for your comment. A chapter has been added to the beginning of the thesis dedicated to the glossary which include an alphabetical list of terms utilized within the work.

   - in the literature review, sections 2.3.1 and 2.3.2 are placed in section 2.3, dedicated to machine learning, although they do not relate to it in any way, so it is more logical to put them in a separate subsection.

Thank you for your comment. I totally agree with you. The comment has been addressed and these two subsections are stated as separate sections.

- on page 55 in Figure 5-2 formulas (11) and (15) are used, but such numeration cannot be in the work.
  Thank you for your comment. This comment has been addressed and the figure has been updated.

- the work contains many references to formulas and figures, which are placed in the following sections, which makes reading difficult, for example, on page 31 there is a link to formula (4.45), which is located on page 49, and further notations from this formula are used starting from page 31 without additional explanation, on page 66 there is a reference to Figure 6-2, although the same plot is shown in Figure 5-5. located on the next page, and to which there is no reference at all in the text of the work, a similar situation is on page 75 with Figures 6-6 and 5-13.
  Thank you for pointing out these issues. The comment has been addressed and the citation of these figures in the thesis are modified.

- the title of section 6.2.10 begins with a lowercase letter.
  Thank you for your comment. This comment has been addressed.

- on page 18 the abbreviation HVDC is not described and there is no explanatory link to the source where one could find out the explanation.
  Thank you for your comment. This comment has been addressed.

- on page 19 at the bottom instead of "prefect" there should be "perfect".
  Thank you for your comment. This comment has been addressed.

- on page 23 there is a typo at the bottom "domain,so".
  Thank you for your comment. This comment has been addressed.

- on page 30 the figures numbers are missing at the bottom.
  Thank you for your comment. This comment has been addressed.

- on page 43 in formula (4.17) the elements of the matrix H are not described and have not been introduced previously.
  Thank you for pointing out this issue. The H matrix elements are modified, and the equations provided after the matrix explains the details of calculating those elements.