

Title

Simon dos Reis Spedsbjerg

August 6, 2024

1 Problem

As data collection forever increases and systems ever become more dependent on the quality of the data. Faulty data may damage the quality of analysis that is performed on the data, and outlier data may indicate another system is failing or an uncommon phenomenon has occurred in the sampled group. Anomalies has become an interest in many sectors, Mukherjee, A[2] cite transport, manufacturing, water, and power, making use of IOT-based sensing. This problem is also an interest when it comes to cyberattacks. Gauthama et al. worked on analysing possible attack detection systems for a multi-point IOT system[1].

2 Project

The project's goal is to develop a new system that detects anomalies based on different variables dependencies on each other. As a variable value changes, its change may be a reflection of an error on another variable. As a web of variables all connected, the program should be able to pinpoint the problem area based on the afflicted variables, the degree of affliction, and the spread of the affliction. The project's goal is to develop the product and compare it to similar solutions in the aspects of speed, accuracy, and precision through relevant analysis methods. Determining the pros and cons of this system or any algorithms developed for it compared to state-of-the-art systems or algorithms may further possible future solutions for software problems.

There are two types of anomalies that the system intends to handle: the normal anomaly, which is a one-off anomaly that is rare or just a misreading, and the discord anomaly, which is a recurring anomaly that can indicate a fault, failure, or special event at some point in the observed systems. By the different sensor variables having relations among one another, it is hoped that discords can be determined with high accuracy and precision.

2.1 Expected Challenges

throughout the project I expect that I would have to find solutions for the following challenges.

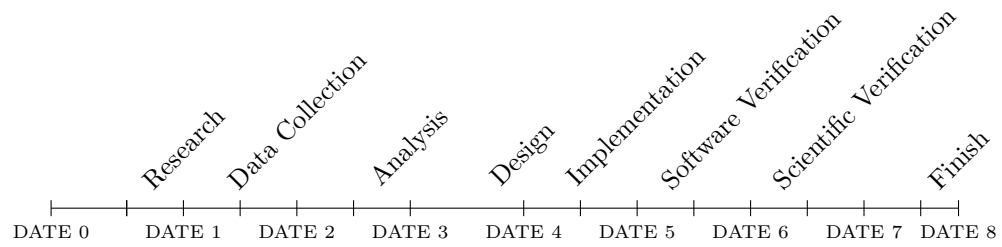
- Making the system be able to differentiate between an anomaly and a discord with a small sample size.
- Determining whether to focus on supervised learning, unsupervised learning or both.
- Availability of data which has the quality required for this project. Either it is successful to gain such data, or a virtual twin of machines which this system would function with has to be made.

2.2 Phases

We intend to go into 7 phases during this project.

- Research, develop a clear and thorough understanding of the current state of research in anomaly detection and fault detection and diagnosis by creating a systematic literature review of the subjects.
- Data Collection, Collect or create relevant data depending on the results of the Research which is relevant to the subject. As the results will be drawn on the solutions ability to handle the data, the quality of the data has a high impact on the project.
- Analysis, analyse the problem and with the help of software engineering methods based on the Research results.
- Design, design a theoretical solution to the problem specified in the Analysis with software engineering methods.
- Implementation, develop the project based on the results of the Analysis and Design using software engineering methods.
- Software Verification, verify that the solution successfully conforms to the requirements determined in the Analysis and implements all critical features.
- Software Verification, verify the results from the project and check if it fully fills the hypothesis. From this draw a conclusion.

In the following, we see the expected timeline for the project.



References

- [1] “AICrit: A unified framework for real-time anomaly detection in water treatment plants”. In: *Journal of Information Security and Applications* 64 (2022), p. 103046. ISSN: 2214-2126.
- [2] “Unsupervised constrained discord detection in IoT-based online crane monitoring”. In: *Advanced Engineering Informatics* 60 (2024), p. 102444. ISSN: 1474-0346.