Title: A Meta-Analysis of the Anomaly Detection Problem

Authors:

- Andrew Emmott, Oregon State University

- Shubhomoy Das, Oregon State University

- Thomas Dietterich, Oregon State University

- Alan Fern, Oregon State University

- Weng-Keen Wong, Oregon State University

Abstract:

The paper presents a comprehensive meta-analysis of the anomaly detection problem. It begins by discussing various approaches to benchmarking anomaly detection algorithms and creating a diverse corpus of anomaly detection benchmarks. The benchmarks are designed to reflect real-world applications by considering factors such as point difficulty, relative frequency of anomalies, clusteredness of anomalies, and relevance of features. Multiple anomaly detection algorithms are applied to the benchmark corpus, generating a large collection of experimental results. The paper analyzes these results to gain insights into experimental design, evaluates the performance using different metrics, and provides guidelines for future experimental design. Additionally, it compares the results against a trivial solution to normalize the reported performance. The article aims to contribute by providing a publicly-available corpus of anomaly detection benchmarks, an ontology for describing anomaly detection contexts, a methodology for benchmark creation, and a discussion on measuring success in the field.

Introduction:

The introduction highlights the importance of anomaly detection in various domains, such as computer security, astronomy, environmental monitoring, machine component failure detection, and cancer detection. It emphasizes the lack of a standard methodology for evaluating anomaly detection algorithms, leading to difficulties in comparing different approaches and understanding the factors influencing algorithm performance. The paper aims to address these issues by conducting a meta-analysis of the anomaly detection problem.

Methodology:

The paper presents a standardized evaluation methodology for statistical anomaly detection, building upon previous work. It describes the refinement of the anomaly detection problem definition and reviews existing approaches to evaluate anomaly detection methods. The authors define requirements for experimental methodology and identify four relevant problem dimensions. They detail the benchmarking methodology, including procedures for meeting the requirements and controlling the problem dimensions. The study utilizes a set of anomaly detection algorithms, and their parameterization details are provided in an appendix.

Results and Discussion:

The paper identifies evaluation metrics, such as the area under the ROC curve (AUC) and Average Precision (AP), and describes the statistical hypothesis tests applied to each algorithm's results on each benchmark. It summarizes the hypothesis tests and discusses the results, comparing measurable quantities against their control groups. Alternative views of the results are presented to highlight algorithm properties and the challenges of comparing algorithms in the anomaly detection field. The findings are discussed globally, providing recommendations for future work in the field.

Conclusion:

The conclusion summarizes the contributions of the paper, which include providing a publicly-available corpus of anomaly detection benchmarks, an ontology for describing anomaly detection contexts, a methodology for benchmark creation, guidelines for experimental design, and insights into evaluating success in the field of anomaly detection.

Note: The content provided is a brief overview of the paper "A Meta-Analysis of the Anomaly Detection Problem." The full content of the paper may contain additional details, experimental procedures, and findings that are not included in this summary.