

Review

of the paper entitled

*A Novel Change Detection Approach on Content Delivery Network
(CDN) using Data Streaming Analytics*

submitted by

Indra Awal Priyanto, Andras Wolosz, Zoltan Kadar, Adam Baba

Reviewer: *Karcag Tamas*

May 7, 2023

The article subject to this review task is a writing practice with imaginary results, but actual literary research for the Research methodology class at ELTE Faculty of Informatics.

Recommendation

strong accept / **accept** / weak accept / borderline / weak reject / reject / strong reject

Summary of the Content, Evaluation

The article discusses detecting changes in data streams, particularly in content delivery networks (CDN), and proposes a novel detection method that uses different statistical modes to measure its effectiveness. The statistical analysis provides insight into the performance, sensitivity, and specificity of various machine learning models and balancing techniques. The article also introduces a new approach to handling stream data using a sliding window and shows its efficacy in improving results.

The study was conducted on an imbalanced dataset, and the V-GAN model was found to perform significantly better than the other models. In both static and streaming contexts, the random forest (RF) model outperformed the other machine learning techniques.

Technical Content and Accuracy

(5 points)

The technical content appears to be of high quality and accuracy, and the references have been properly utilized.

Significance of the Work *(4 points)*

The field of data streams and their associated services is rapidly expanding in the contemporary internet landscape, with an increasing number of such features being developed. While the current solution aims to enhance the quality of service, the utilization of virtual geo-separation results in lower costs for data stream providers.

Appropriate Title, Introduction and Conclusion *(5 points)*

The title is clear and concise, conveying the topic effectively. The introduction and conclusion sections of the document are well-crafted, providing comprehensive and precise information.

Overall Organization *(4 points)*

The document exhibits a well-organized structure, containing all the necessary and mandatory sections. The ordering of the sections is appropriately precise. Additionally, the length ratios of the different sections are correct, striking a balance between providing comprehensive coverage of the subject matter.

The section on related works lacks coherence, with disjointed paragraphs that fail to establish clear connections between them.

Appropriateness for the Conf/Journal *(5 points)*

The research paper is entirely suitable and appropriate for presentation at the conference.

Style and Clarity of the Paper *(5 points)*

The writing style employed in the paper is straightforward and easy to read, with clear and concise language used throughout. The text is precise and unambiguous, making it readily understandable to a broad range of readers.

Originality of the Content *(4 points)*

The paper presents a method that is unquestionably beyond the state-of-the-art, with extensive evidence of its efficacy. Moreover, the references have been accurately and effectively utilized to support the claims made in the paper.

Sum Score *(32 points)*

Novelty of Results

The paper presents a method that has been thoroughly tested and is based on cutting-edge solutions. The approach utilizes modern machine learning models and balancing techniques to optimize data stream handling and enhance the quality of service and the quality of experience. In addition, a sliding window technique is employed to improve the combination of accurate data stream handling and geo-supported services via Spark are also available and used.

The solution is predominantly founded on Spark's functionality, which is a contemporary geo-distributed big-data service equipped with advanced data analysis and processing tools.

Appropriateness of the Methods, Validation, References

The methods employed in the study demonstrate a high degree of appropriateness, with references and papers carefully selected from reputable and trustworthy sources and authors (IMO). The validation process conducted to evaluate the efficacy of the methods is comprehensive and detailed, ensuring that the results are both reliable and accurate. Overall, the study appears to be conducted with a high degree of rigor and attention to detail.

Comments on Errors, Typos, Grammar, Figures

Some figures are in incorrect sections:

- Fig 10. are referenced in section IV.A but the figure is in IV.B

The term 'hence' is frequently utilized in the research paper. In the Results section, the word appears as the first word in two or more adjacent sentences.

Typos or spell checkings:

- In I.D the 'imabalanced' word should be 'imbalanced'
- In III.A the 'nformation' word should be 'information'
- In III.D the 'environement' word should be 'environment'

Proposal for Improvements

The objective is to improve the readability and clarity of the Related Works section by simplifying the language and making it more user-friendly. The current overview of the section appears convoluted and difficult to navigate, making it challenging for readers to fully comprehend the presented information. Additionally, identifying the state-of-the-art tools and methods used in the study can be a time-consuming and daunting task.

Reviewer's Confidence

expert / high / **medium** / low / none