

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: *Kristóf Bezzegh*

May 9, 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 the challenges of detecting and quantifying changes in data streams, particularly in a CDN environment. A two-window paradigm is proposed for change detection and algorithms are developed with proven guarantees on their sensitivity to change, robustness against false alarms, and running time. The article aims to obtain reliable change detection and a comprehensible description of the nature of the change. The authors conducted experiments on an imbalanced CDN dataset using various balancing techniques and evaluated the performance of different machine learning models using AUC, sensitivity, and specificity scores. MLP was the best in terms of AUC, but RF was chosen due to its lower model complexity and numerical superiority. T-tests were conducted to compare the statistical significance of the models. In the streaming context, a sliding window approach was used, and RF outperformed other models.

Technical Content and Accuracy *(5 points)*

The technical content and accuracy of the research paper was impressive, with thorough analysis and accurate data presented.

Significance of the Work *(5 points)*

The paper identifies a significant difference between processing data streams and static data, which is that data streams have a temporal dimension, and the underlying process generating the data stream can change over time. The paper proposes a change detection algorithm based on a two-window paradigm that can detect when the distribution of data changes and quantify and describe the change. The paper also discusses the statistical requirements of sudden changes and their applications in Quality of Service and Data Mining. The proposed change detection algorithm can help detect sudden changes within a network, enabling better management of the continuous stream of data in CDN data.

Appropriate Title, Introduction and Conclusion *(4 points)*

The title is clearly and concisely captures the essence of the study and draws the reader in with its captivating and thought-provoking nature. The introduction is well-written and provides a clear and concise overview of the topic, setting the stage for a comprehensive investigation. The study presents a promising approach to identifying sudden change-points in CDN data, but further research is needed to address the limitations and validate the proposed methods in real-world scenarios.

Overall Organization *(5 points)*

The organization of this research paper is exceptional, as it starts with a clear and concise introduction that sets the stage for the research, followed by a comprehensive literature review that highlights the most relevant previous works in the field, and then a well-structured methodology section that outlines the data collection and analysis processes used to obtain the results presented in the subsequent results and discussion sections, which are presented in a logical and easy-to-follow manner, and finally a conclusion that effectively summarizes the main findings and their implications for future research.

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

The research paper is an excellent fit for the selected conference/journal, as it provides a novel and significant contribution to the field and presents the results in a clear and concise manner, making it accessible to a wide range of readers.

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

The paper is written in a clear and concise style, with each section flowing seamlessly into the next, making it easy to understand and follow the authors' line of reasoning. The use of visual aids and graphs further enhances the clarity of the paper, making it a pleasure to read.

Originality of the Content *(5 points)*

This research paper presents a unique and innovative approach to solving a problem in the field, offering fresh insights and potential for future developments.

Sum Score *(34 points)*

Novelty of Results

Beyond State of the Art (Related Work)

The paper discusses various techniques and algorithms used for sudden change-point detection in CDN data, including data mining algorithms, online classification and clustering, outlier detection, and association rule mining. One of the novelties of the paper is that it provides a comprehensive overview of the research done in this field, including the challenges involved in sudden change-point detection, such as the presence of missing values in the data, and the need for scalable systems. The paper also discusses recent developments in the use of deep learning models and ensemble techniques for sudden change-point detection in real-time data streams, which can help address the overfitting issues associated with previous studies. Overall, the paper presents a valuable contribution to the research on sudden change-point detection in CDN data and provides useful insights into the various techniques and challenges involved in this field.

Appropriateness of the Methods, Validation, References

The methods used in this research paper are highly appropriate and well-suited for the study. The researchers have carefully selected the most cutting-edge and innovative techniques in the field, and have applied them in a rigorous and systematic manner. The results are therefore both reliable and insightful, and represent a significant contribution to the advancement of knowledge in this area. The validation of the study was thoroughly conducted, and the references cited were well-researched and reliable, adding credibility to the findings.

Comments on Errors, Typos, Grammar, Figures

A paper with some typos can still have valuable insights and contribute to the field, but it is important for the author to proofread and correct the errors to ensure the paper's credibility and professionalism.

Typos:

- In I.D the 'imabalanced' word should be 'imbalance'
- In III.A the 'nformation' word should be 'information'

Proposal for Improvements

One potential improvement for the paper would be to revise the articulation to make it more clear and concise, while maintaining the engaging writing style that made the paper enjoyable to read. Additionally, including more subheadings and a clear table of contents could make it easier for readers to navigate through the content.

Reviewer's Confidence

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