

# Optimized plagiarism detection is 400x faster than conventional methods

September 28 2022, by David Bradley

---



Credit: Unsplash/CC0 Public Domain

In a world where so much information is so readily available to students, educators and student assessors must constantly fight against plagiarism. The time and effort required by an examiner potentially faced with

hundreds of essays to check for such problems however small is huge. Semi-automated tools exist for identifying plagiarism in a sample of text but these too take up computing resources and are often unwieldy and more suited to single documents.

Writing in the *International Journal of Innovative Computing and Applications*, a team from Australia and Sri Lanka has developed a [new computational approach](#) to plagiarism detection that uses vector space and exploits the architecture of graphics processing units and their compute unified device architecture (CUDA) rather than a conventional computer chip, a [central processing unit](#), CPU.

Jiffriya Mohamed Abdul Cader of the Sri Lanka Institute of Advanced Technological Education Sammanthurai, Akmal Jahan Mohamed Abdul Cader of the South Eastern University of Sri Lanka, Hasindu Gamaarachchi of the University of New South Wales, Australia, and Roshan G. Ragel Faculty of Engineering, University of Peradeniya, Sri Lanka explain that conventional serial testing of 1,000 documents can take half an hour.

The prototype of their GPU approach improves on that significantly, taking just 36 seconds to process the same dataset and flag any plagiarized sections of text. However, the researchers further optimized their prototype and were able to reduce processing time to just 4 seconds for one thousand documents. That's almost 400 times faster than conventional approaches. Such speed would be a boon to examiners faced with hundreds if not thousands of student-submitted documents to check for [plagiarism](#).

The next step will be to test the same approach on text found in other kinds of document rather than simply straight-text essays, including notebooks, assignments, reports, theses, and such.

**More information:** Jiffriya Mohamed Abdul Cader et al, Optimization of Plagiarism Detection using Vector Space Model on CUDA Architecture, *International Journal of Innovative Computing and Applications* (2021). [DOI: 10.1504/IJICA.2022.10042480](https://doi.org/10.1504/IJICA.2022.10042480)

Provided by Inderscience

Citation: Optimized plagiarism detection is 400x faster than conventional methods (2022, September 28) retrieved 18 April 2024 from <https://techxplore.com/news/2022-09-optimized-plagiarism-400x-faster-conventional.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.