PROTOTYPE TOOL FOR MISCONDUCT DETECTION: A DIGITAL FORENSICS APPROACH

Clare Johnson[1], Ross Davies[2], Mike Reddy[2]

[1]*University of South Wales, United Kingdom*
[2]*University of South Wales, Newport Campus, United Kingdom*

Keywords

Plagiarism, contract cheating, academic misconduct, academic integrity, detection, OOXML, forensics

Abstract

Digital forensics techniques are being used more and more frequently to gather evidence in criminal investigations, particularly those relating to cyber crime or crimes relating to Intellectual Property. However, these techniques are not yet widely used in determining the authenticity of student submissions, despite the significant problems of plagiarism and contract cheating in academia. Current methods of misconduct detection focus on text-matching software programmes which identify text that matches, or is very similar to, existing digital work. Some of these text-matching software providers have recently rolled out authorship tools that provide basic metadata about the submission. This may include analysis of the language and writing style of the author and compare these findings across cohorts as well as across previous submissions by the same student. The information provided by these additional reports has been shown to have a positive effect on assessors' ability to detect contract cheating (Dawson, Sutherland-Smith & Ricksen, 2019).

Plagiarism detection can be described as both extrinsic and intrinsic (PAN, n.d.). Extrinsic detection compares the document to existing work in order to evidence plagiarism (e.g. text-matching) and intrinsic plagiarism detection analyses the input document using 'stylometry to examine linguistic features of the document' (Foltýnek, Meuschke & Gipp, 2020), detecting different writing styles within a single document and identifying features specific to certain authors, similar to the authorship tools already mentioned. Neither extrinsic nor intrinsic detection methods consider the document as an object in its own right, or analyse the information that is available *behind* the text that is seen in print or on screen.

Word documents are constructed using Open Office Extensible Mark Up Language (OOXML) format. During the writing process every piece of text is automatically allocated an edit run value (rsidR) by the software. Text written in one editing session (i.e. before a document save, whether manual or automatic) shares the same rsidR value. Text that is edited after being written introduces a new rsidR value as do additions and style changes. Analysing the rsidR values and providing a visual output of the editing can provide valuable insight into how the document has been written.

This paper presents a prototype tool for a novel approach to plagiarism, collusion and contract cheating detection, building on previous experiments by the authors. 'Clarify' (working title) extracts the metadata and forensic artefacts from work submitted in Word format, looking behind the text itself for detailed information on how the document has been written as described above. A file is passed into the software and automatically decompressed into its component parts, and the various artefacts are displayed in an easy to read report, including: visual display of edit runs, number of edit runs, list of edit run values, number of revisions, date created, author, total time spent editing, number of font changes, number of format changes, number of font size changes, evidence of white text and uncropped images. rsidR values are counted and anomalies flagged.

Experiments were initially carried out on submissions that had already gone through a misconduct panel, and these are now continuing on authentic work in an attempt to build a benchmark of what an authentic assessment should look like. A prototype flagging system is being created so that unusual values, which are significantly higher or lower than those that would be expected from an authentic document are highlighted in the report. For example, it could be expected that the number of rsidR values will correlate with the length of the document, a longer document having more unique rsidR's than a shorter one. Documents that fall outside of what is deemed 'normal' will be highlighted in the report. Similarly, a file creation date that precedes the assignment release could suggest that the student has reworked a previous student's submission.

Dawson, Sutherland-Smith & Ricksen (2019) note that the availability of software tools (authorship) support and improve detection of misconduct, not least by simplifying the detection process, but also by providing objective evidence for misconduct panels, as well as raising awareness that these types of misconduct activities exist (thus making the assessor more alert to them). Whilst 'Clarify' itself will not provide a silver bullet for solving misconduct, it could, in time, be an excellent addition to the techniques used by software such as Turnitin, sitting alongside authorship tools as yet another layer of detection. Of course, all tools require a good degree of care when interpreting the results, but early indications suggest that despite being in its infancy, the application of digital forensic tools as provided in this proof-of-concept could provide a very useful additional tool for academics to use when assessing student submissions for authenticity, drawing attention to anomalies in an easy to digest format that could greatly both speed up the process of detection as well as improve detection rates.

References

Dawson, P., Sutherland-Smith, W., & Ricksen, M. (2019), Can software improve marker accuracy at detecting contract cheating? A pilot study of the Turnitin authorship investigate alpha, *Assessment and evaluation in higher education*, pp. 1-10, https://doi.org/10.1080/02602938.2019.1662884

Foltýnek, T., Meuschke, N., & Gipp, B. (2020). Academic Plagiarism Detection: A Systematic Literature Review, *ACM Computing Surveys*, *52*(6) Article 112, pp. 1-42. https://doi.org/10.1145/3345317

PAN (n.d.). *Shared Tasks.* https://pan.webis.de/shared-tasks.html#intrinsic-plagiarism-detection