**Construction plagiarism findings in scientific research (hùng vương univercity)**

The research was developed in 2017 by two lecturers of Hung Vuong University, Dinh Thai Son and Nguyen Thi Hao.

Research methods:

- detect copy according to PAN solution.

Copy detection model integrates with single value separation algorithm.

research results:

TABLE 1. Create test documents

|  |  |  |
| --- | --- | --- |
| How to create documents to check | Rate of change from the source material | Document number suspected |
| Copy and paste | 0% | 580 |
| Copy and paste with little modifications | 10-15% | 580 |
| Copy and paste have major changes | 30-45% | 579 |
| Total | 0-45% | 1739 |

TABLE 2. Results of measuring PAN rating

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data set | Plagdet | Precision | Recall | Granurality |
| Copy and paste | 0.9639 | 0.9355 | 0.9904 | 1.0000 |
| Copy and paste with little modifications | 0.9189 | 0.9138 | 0.9319 | 1.0057 |
| Copy and paste have major changes | 0.7961 | 0.8958 | 0.7359 | 1.0264 |
| Total | 0.8951 | 0.9151 | 0.8886 | 1.0101 |

To test the system in accuracy factor, we created some documents to check from 145 source documents (data taken from graduation thesis of Hung Vuong University students).

Table 1 shows how to generate documents to test and the number of documents to check. The test results on the data set are shown in Table 2.

We measured 4 PAN rating factors that were used to evaluate a plagiarism detection system in PAN (Kasprzak and Brandejs, 2010). These include plagdet, precision, recall, and granularity. In general, with SVD algorithm applied, it shows that the value asymptotic to the algorithm on PAN used.

**An introduction to anti-plagiarism software**

Copyscape was launched in 2004 by Indigo Stream Technologies, Ltd., co-founded in 2003 by Gideon Greenspan.

Providing the URL or the text of the original content, Copyscape returns a list of webpages that contain text that is similar to all or parts of this content. It also displays the highlighted relevant text on the found web page. Copyscape banners can be placed on a website to warn potential plagiarists not to steal content. Copysentry monitors the web and sends email notifications when new copies are found, and Copyscape Premium verifies the originality of content purchased by online content publishers.

Copyscape uses the Google Web API to assist with its searches. Copyscape uses a set of algorithms to identify copied content that has been modified from its original form.