Abstract

In today word copying something from other sources and claiming it as an own contribution is a crime. We have also seen it is major problem in academic where students of UG, PG or even at PhD level copying some part of original documents and publishing on own name without taking proper permission from author or developer. Many software tools in exist to find out and assist the monotonous and time consuming task of tracing plagiarism, because identifying the owner of that whole text is practically difficult and impossible for markers. In our presentation we have focused on practical assignments (projects) as well as written document which is to be submitted by students in to college or university. Because of this crucial task and day by day increasing research in different fields, industry, academy people demanding such software to detect whether submitted articles, books, national or international papers are genuine or not. In this paper, our algorithm divides submitted articles in small pieces and scans it to compare with connected databases to the server on internet. Some existing work compares submitted articles with previously submitted articles i.e. with existing database.
- Tommy W. S. Chow and M. K. M. Rahman has developed a approach of Multilayer SOM With Tree-Structured Data for Efficient Document Retrieval and Plagiarism Detection; IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 20, NO. 9, SEPTEMBER 2009.
- S. Schleimer, D. Wilkerson, and A. Aiken; Winnowing: Local algorithms for document fingerprinting; in Proc. 22nd Association for Computing Machinery Special Interest Group Management of Data Int. Conf., San Diego, CA, Jun. 2003, pp. 76–85.
- Xin Chen, Brent Francia, Ming Li, Member, IEEE, Brian McKinnon, and Amit Seker Shared Information and Program Plagiarism Detection; IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 50, NO. 7, JULY 2004.
- Chao Liu, Chen Chen, Jiawei Han, Philip S. Yu; GPLAG: Detection of Software Plagiarism by Program Dependence Graph Analysis; KDD'06, Philadelphia, Pennsylvania, USA. August 20–23, 2006.

**Index Terms**

Computer Science 

Algorithms

**Keywords**

Document retrieval; Plagiarism 

Algorithm Karp-Rabin; plagiarism detection String matching. )