Source code plagiarism is becoming a common practice among higher education community. People duplicate and modify the source code of other people and show the program as their own program. In this paper, we want to draw researchers’ attention towards this problem and projected a novel approach which detects plagiarism in C language code by converting it into assembly language which is done with the help of GCC compiler. Assembly language converted by the compiler is not sensitive to all type of different code transformation, for example-swapping variable names, reformation of language, adding extra comment or blanks. Therefore, assembly language gives rise to reduced amount of variations, if there is a modification in the original code. Previous works in plagiarism compares the whole program but in this paper, we proposed a method which split the C program into assembly language code and divide each function of program into blocks and blocks are transformed into token strings. This method compares each function with other program function and provides a statistical output, according to the token string likeness of that function. If the output is above assigned specific plagiarism similarity threshold value then it counts under the case of plagiarism.
References

governance/policies-and-documents/statement-on-plagiarism.
3. Ottenstein, K.J.: An Algorithmic Approach to the Detection and Prevention of
4. Schleimer, S., Wilkerson, D., Aiken,
5. Wise, M.J.: YAP3: improved detection of similarities in computer program and other
texts. In: Proceedings of the Twenty-Seventh SIGCSE Technical Symposium on
Computer Science Education, vol. 28(1), pp. 130–134. Association for Computing Machinery,
New York (1996)
7. H. L. Berghel and D. L. Sallach. Measurements of program similarity in identical task
1981
9. M. J. Wise, “Detection of Similarities in Student Programs: YAP“ing may be Preferable to
programs,” Fakultät für Informatik Technical Report 2000-1, Universität Karlsruhe, Karlsruhe,
2012
11. Sam Grier. A tool that detects plagiarism in Pascal programs. ACM SIGSCE Bulletin
12. DickGrune website regarding to similarity measure
URL:http://www.dickgrune.com/Programs/similarity_tester/
13. Jplag tool site URL: http://jplag.ipd.kit.edu
“Software Plagiarism Detection Techniques: A Comparative Study”, International Journal of
Computer Science and Information Technologies, Vol. 5 (4), ISSN: 0975-9646
15. Enrique Flores, Alberto Barrón-Cedeño, Paolo Rosso, Lidia Moreno, Jun 2012,
“DeSoCoRe: Detecting Source Code Re-Use across Programming Languages”, NAACL-HLT
2012
16. Shan S., Guo F., Ren J.: similarity detection method based on assembly language and
string matching
17. Gupta A., Singh S.: lexical analysis for the measurement of conceptual duplicity between
C programs, in proceedings of vol. 1 issue, AUGUST 2013.
code datasets, in proceeding of the computer journal (November 2005)

**Index Terms**

Computer Science  
Information Sciences

**Keywords**

Plagiarism, assembly language, string similarity, Plagiarism detection method, token string