Advances and Applications in Mathematical Sciences
Volume 20, Issue 3, January 2021, Pages 391-396

INTEGRATED CLOUD COMPILER WITH PLAGIARISM CHECKER

NIDHI VARDHAN, APURVA SHARMA and VIKASH YADAV

Computer Science and Engineering Department
ABES Engineering College
Ghaziabad, U.P India
E-mail: nidhi.16bcs2007@abes.ac.in
apurva.16bcs1196@abes.ac.in
vikash.yadav@abes.ac.in

Abstract

The Internet has changed so much in our lives. People now rely heavily on the Internet for their activities. Therefore, we have created an online compiler called Internet Compiler. The main aim of this project is to write a Java program, compile and debug it online and help prevent exploitation. This frees up the overhead of using Java development kits and hardware storage on every client machine. They are only associated with the server and the server Java compiler. Therefore, it runs Java programs and displays an error message when an error occurs on the client machine. These are the Online Compiler Service (SASS) where the download and installation of a separate compiler on each machine stops. Code errors and outputs are stored in the cloud storage.

I. Introduction

A compiler is a special program that converts a code statement written in high level language into a machine language. The main objective of the compiler is to change the code written in a language without changing the meaning of the program. The compiler needs to be installed on the client's machine so it will require machine storage. To solve the compiler's storage and portability problems, we are using cloud computing. Users without installing any compiler require submitting the program to the user interface provided. The compiler will compile and run the server program. The output

2010 Mathematics Subject Classification: 68N20, 68P20.
Keywords: Cloud Computing, Compiler, Plagiarism Checker.
Received May 20, 2020; Accepted July 31, 2020

is then returned to the user. Plagiarism is copying another person's code and pretending that they are their own code. To detect plagiarism in the code, plagiarism checker is used. Plagiarism checker is software that can be used to cross-check codes for duplicate content.

Objective: The main objective of the project is to write a Java program, compile and debug it online and also helps prevent plagiarism. Cloud Compiler can reduce storage space and movability problem by using cloud computing which helps save time, storage and cost. It is the most appropriate tool for compiling code, removing errors, and debugging it. The cloud compiler will be integrated with plagiarism checker to reduce the possibility of theft.

Problem Identification and Definition: A compiler that has to be manually installed on every system requires a lot of machine storage and also needs to be configured when not installed using the default parameters. If the use of a single system is not allowed, it is difficult to move the same source code across different machines. Another consequence is that we have to set up a separate complement on each language we want to work on. The plagiarism checker is integrated into the source code to reduce the possibility of plagiarism.

II. Literature Survey

In addition to the implementation viewpoint in [1], the focus is on how instructors can manually detect exploitation without comparing source files. Common Indicators are Output, Behavior, Structuring, Coding-Styles, and Flipping. Pre-processing of the source file reduces the source code to four strings, namely file structure, identifiers, literals, and tokens. The preprocessor consists of two scanners. The first scanner is line-oriented and the second scanner is a standard C-scanner, which eliminates whitespaces and comments, but does not include them. The De Lex-Scanner Generator is used to create table driven finite state automata from regular expressions sets.

The four string representations are compared to the corresponding string representations stored in the repository. Except for a list of identifiers, the greedy-string-tiling algorithm is used. The algorithm divides the strings into

character-lines of minimum match length and tries to match these sequences. Successive matches are checked until a different character appears. The algorithm returns the percentage of characters that match the total number of characters in the string.

Data mining in the paper [2] can help detecting plagiarism and improve the efficiency of the code. This paper explains various forms of plagiarism which may be recognizable or may be complicated. Some forms of coping and paste, reuse of existing work, self-exploitation.

The algorithms used are the longest common row word algorithm, MDR (Match Detect Reveal).

In [3] it states that cloud computing uses resources online. Online Service is used to solve problem of hard disk storage. Software-As-A-Service Cloud Computing is used. SaaS is a software service that allows people to sell, buy, build and use software. Cloud providers, who use scalable IT services and computing resources, use the public cloud resource to create a virtual cloud. Software as a service uses service oriented architecture where applications can easily communicate with them.

Compiler: A compiler is a special computer program that is used primarily to convert source code from programming language to object code.

Results:

Compilation time: The given source code must be converted to object code for the executable program.

Absolute Running Time: The absolute time taken to process a set of codes.

CPU time: The time used by the central processing unit to perform the instructions is specified as CPU time.

Memory Peak: This is a memory that the CPU can access quickly, without using hard drive memory.

Absolute service time: Service time is the time taken to send and receive a response for a set of signals.

Output: The result that arises after compiling.

In [4], some tools are described that can help identify possible exploitation in an assignment. Therefore, the only solution that can be laborious and error-prone is through the handchecking of documents having similarities.

Paul Heckel's algorithm: This algorithm distinguishes between changed lines, which help to deal with independent advertising order.

It is important to have knowledge of the various forms of plagiarism as discussed in [5]. It prevents severity and damages. This paper explains the various methods based on machine learning techniques to detect the plagiarism.

In [6], security to the editor is provided by using RSA Algorithm. It is very efficient to encrypt and decrypt the files. Types and characteristics of compiler are also defined. It also explains the advantages and disadvantages of Online Compiler.

Cloud computing main qualifying technology is virtualization defined in [7]. It enables hardware based devices to be electronically separated into one or more virtual devices which will help in completing the tasks efficiently. SaaS, Paas, IaaS are various services of cloud computing.

III. Proposed Methodology

The proposed work is carried out in following steps:

Step 1. Gaining efficient knowledge about cloud computing (Seminars, research paper, books, conference) and all the technologies that are going to be used in website development i.e. Information Collection.

Step 2. Design and development of compiler:

Backend:

Servlet: Used to make the web pages dynamically

Ajax: Ajax is asynchronous Javascript + XML. It is used to make the web page interactive.

Frontend: HTML, CSS, Bootstrap

Step 3. Pre processing of the data is carried out followed by text classification.

Plagiarism Checker: Data mining techniques are used.

Step 4. Texts in the source code are classified on the basis of their features.

Step 5. Clustering, Information extraction and summarization of the data goes on in cycle.

Step 6. All similarities are matched and the plagiarized data is clustered.

Cloud Compiler integrates with plagiarism checker.

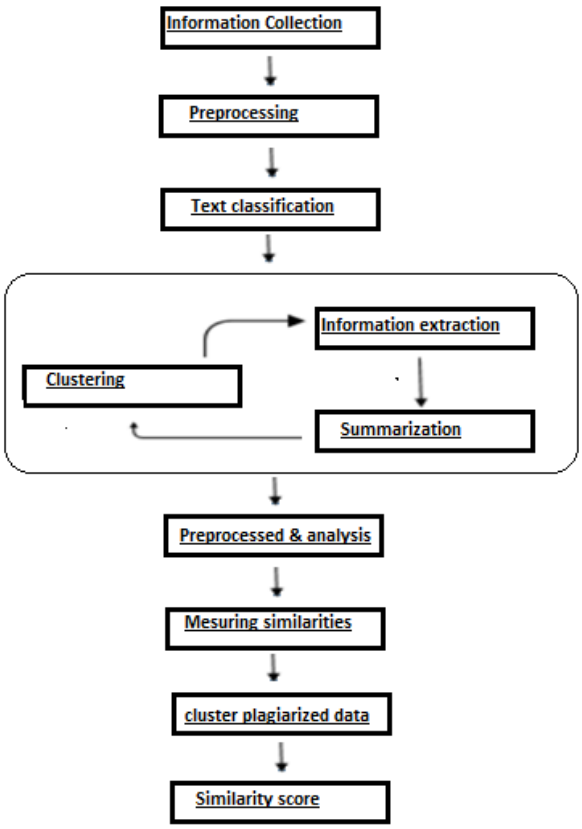

Figure 1. Proposed Method.

IV. Implementation and Results

The implementation of the work is depicted with the following figure.

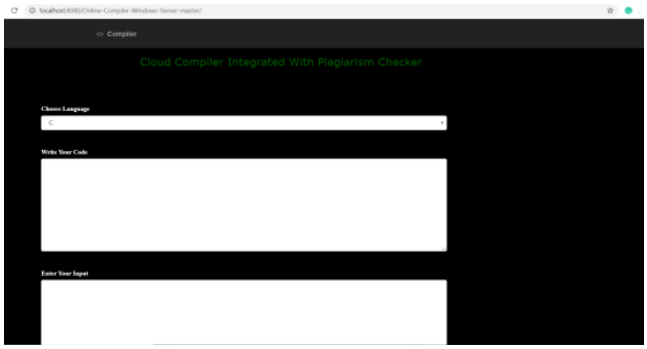

Figure 2. Snapshot of Compiler.

V. Conclusion

The proposed work presents an online compiler which compiles the source code of languages like C and Java. The work has done in Java. Addition of cloud services enhances the functionality of this online compiler which makes it more useful.

References

[1] Online detection of source-code plagiarism in undergraduate programming courses D. Pawelczak Faculty of Electrical Engineering and Computer Science University of Bundeswehr Munich (UniBw M), Neubiberg, Germany.

[2] Plagiarism Detection Process using Data Mining Techniques, Mahwish Abid, Muhammad Usman, Muhammad Waleed Ashraf.

[3] Online Compiler with Plagiarism Checker, G. V. Aarthi, Abishek Rajagopal, Mukundhan Lakshmanan, Kowshik kumar Arulprakasam.

[4] Cloud Based Compiler Sajid Abdulla, Srinivasan Iyer and S. I. E. S. Sanjay Kutty, Graduate School of Technology, Nerul, Navi Mumbai, Maharashtra, India.

[5] Plagiarism: Taxonomy, Tools and Detection Techniques, Hussain A Chowdhury and Dhruba K Bhattacharyya Dept. of CSE, Tezpur University.

[6] Review Paper on Online Java Compiler, Shamali Kokare, Divya Chauhan, Jyoti Mishra and Aarti Sakore, Prof. Manisha Singh Shamali, Computer Science Department of Engineering, Dhole Patil College of Engineering, Maharashtra, India

[7] ONLINE JAVA COMPILER USING CLOUD COMPUTING, Aarushi Verma Namita Garg Dept. Computer Science Dept. Computer Science.

[8] https://www.researchgate.net/publication/276922178_Online_compiler_as_a_cloud_service