

PLAGIARISM CHECKER

Degree Project Report



Raffi Ahmad

01-133082-129

Asim Zahir Abbasi

01-133082-155

SUPERVISOR

MR. RASHID KAREEM SIDDIQUI

DEPARTMENT OF COMPUTER AND SOFTWARE ENGINEERING

BAHRIA UNIVERSITY ISLAMABAD

DEDICATION

Dedicated to our beloved family, faculty and friends.

ACKNOWLEDGEMENTS

First of all we would thank Allah Almighty, The most Beneficial, The most Merciful, who is the ruler and The King of mankind.

We would thank all the teachers of Bahria University Islamabad who have helped us during our time here.

We would also specially like to thank Mr. Rashid Kareem Siddiqui who has supervised this project. This would not have been possible without his guidance and help. He has provided us great assistance in the development of this product.

We would also thank our beloved parents who have invested their money on us and without their support we would not have been able to do this project.

Raffi Ahmad

Asim Zahir Abbasi

ABSTRACT

Plagiarism detection tool is software which can determine the amount of data that has been copied in a particular document from the internet. To detect plagiarism an algorithm is implemented which is a string matching algorithm that is used to find a set of pattern string in a text using hashing. It can rapidly search through a paper for instances of sentences from the source material, ignoring details such as case and punctuation and as there is the abundance of sought strings, single string algorithms are impractical. The key to its performance is the efficient computation of hash values of the successive substrings of the text. One popular and effective rolling hash function treats every substring as a number in some base, the base usually being a large prime.

LIST OF FIGURES

Figure 1.1	Levels of Plagiarism.....	3
Figure 5.1	System Flowchart.....	20
Figure 6.1	Waterfall Model.....	23
Figure 6.5	System Architecture.....	27
Figure 6.6a	Use Case 1: Login.....	28
Figure 6.6b	Use Case 2: File Submission.....	29
Figure 6.6c	Use Case 3: System Response.....	30
Figure 6.6d	Use Case 4: Logging out of the System.....	31
Figure 11.1	Login Window.....	46
Figure 11.2	Home Window.....	47
Figure 11.3	Upload Window.....	48
Figure 11.4	Result Window.....	48
Figure 11.5	Add New User.....	49

LIST OF TABLES

Table 1	String Table.....	35
Table 2	Result Table.....	38

TABLE OF CONTENTS

1	Introduction.....	1
1.1	What is Plagiarism.....	2
1.2	Levels of Plagiarism.....	2
1.3	Objectives.....	3
1.4	Project Significance.....	3
2	Literature Review.....	4
2.1	Types of Plagiarism.....	5
2.2	Related Work.....	6
3	System Analysis.....	7
3.1	Problem Statement.....	8
3.2	System Overview.....	8
3.3	Functions.....	8
3.4	Attributes.....	9
4	Software Requirements Specification (SRS).....	10
4.1	Purpose.....	11
4.2	Scope.....	11
4.3	Perspective and Features.....	11
4.4	Interface Requirements.....	12
4.5	Hardware Requirements.....	12
4.6	Software Requirements.....	13

4.7	Functional Requirements.....	13
4.8	Non-Functional Requirements.....	13
5	Proposed Solution.....	16
5.1	System Description.....	17
5.2	System Flowchart.....	19
6	Design.....	20
6.1	Software Development Life Cycle Model.....	21
6.2	Assumptions.....	23
6.3	Constraints.....	24
6.4	System Environment.....	25
6.5	System Architecture.....	25
6.6	Use Cases.....	27
7	Implementation.....	32
7.1	Algorithm.....	33
7.2	Login.....	34
7.3	File Submission.....	35
7.4	Check Plagiarism.....	37
8	Testing.....	39
8.1	What is testing.....	40
8.2	Testing Methods.....	40
8.3	Specification Testing.....	41

8.4	User Training.....	42
9	Conclusion & Future Enhancements.....	43
10	User Manual.....	45
11	References.....	50