Plagiarisms Bypass on the Go

Pravin Shrinath
Computer Engineering
Department
Mukesh Patel School of
Technology Management and
Engineering
NMIMS University, Mumbai,
India

Saurabh Malgaonkar
Computer Engineering
Department
Mukesh Patel School of
Technology Management and
Engineering
NMIMS University, Mumbai,
India

Yash Radia
Computer Engineering
Department
Mukesh Patel School of
Technology Management and
Engineering
NMIMS University, Mumbai,
India

Nipun Philip
Computer Engineering Department
Mukesh Patel School of Technology Management and Engineering
NMIMS University, Mumbai, India

ABSTRACT
This paper describes a certain way in which a user can bypass the plagiarism check. It focuses on how copied content cannot be detected by a plagiarism checker and the unique content is not affected by much. This idea is implemented as an android application. The paper also includes a case study where a random piece of text is selected; the algorithm is applied and passed though a plagiarism checks. The unique content is still not much affected which is a bad good for those who want to plagiarize.

General Terms
Algorithm, Pattern- Matching, ASCII value

Keywords
Plagiarism, Pattern Matching, Unique Content, Android, Web Programming

1. INTRODUCTION
In this present time, the most common source of information to everyone is the internet. The internet has the answer for almost every question. This can be used in the right as well as in the wrong way. Here arises the issue of plagiarism [1]. Now what is plagiarism? In simple words, plagiarism is just copy and paste of contents from the other sources. Plagiarism is committed when an individual copies and pastes any information from an online or offline source for his/her own specific purposes without even giving a single read or having any knowledge about the information copied. For a student plagiarism [2] can be done in his assignments, in their online/offline exams etc. An online source can be internet sites such as Wikipedia [6] and so on and offline sources are books to the related topic. Now in universities where submissions are done online they run the data given by students through a plagiarism checker. This checker has an algorithm which picks up a large string and compares this string with articles from different sites. This comparison is done by comparing the ASCII values of the strings. The algorithm 'ProPlag' is designed and implemented follows a pattern to change the ASCII value of a few predefined characters of a certain language. Thus due to this the checker gets confused as to which language the text pertains to.

2. DESIGN
2.1 Logical Statement
The search logic in search engines uses basic alphabet structure to retrieve or match the search value. For example if the word Rat is searched on the search engine, and use two different style of R in the word, the word "Rat" is getting completely two different results? Now why this happens is because the English letter "R" is replaced by a very similar looking "R" of another language. Due to this the search engine algorithm gets confused as to which language the word it pertains to. This similar logic can be done with most of the alphabets of English language which leads to the increase in the unique content percentage shown by the plagiarism checker; since the checker has almost similar algorithm to any search engine.

2.2 'ProPlag' Algorithm
Step1: Receive the copied content from the user.

Step2: Scan the full content. [Parser Based]

Step3: Check if (char_existing==char_replacable) [Parser Based] [This condition stands true only if plagiarism is present]

   If yes then goto replace the character with the respective similar character (most common characteristics) from the stored source of replaceable characters.

    Check if all the characters replaceable are modified

     If yes goto step4
     Else goto step3
    Else goto step4

Step4: Construct the new modified content

Step5: End

Fig 1: Example of replacing letter 'R' with similar a character

2.3 System Architecture

The user copies the text and pastes it in the android app along with the other necessary parameters and sends it to the remote server for processing. Then this text is passed through a plagiarism check to see the unique content percentage. Then the required text is then edited by the algorithm and then sent back to the user. A mail is also sent to the user with the edited content and notification of the entire operation.
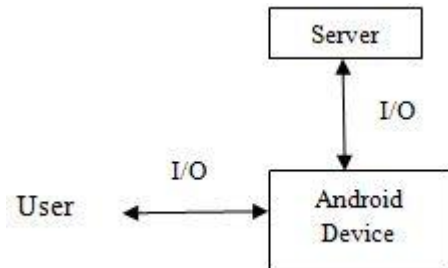

Fig 2: Basic Modular Structure

3. IMPLEM ENTATION

Basically, it is a Web application so it needs a website to be developed which requires web languages such as HTML5 [3], CSS, JavaScript and a little bit of PHP [4] (server side scripting) for mailing purpose. The Android app was made in Eclipse [5] software. For the app to run smoothly the client only needs an android phone with android [7] version above 2.2 with active internet or WIFI connection with a minimum bandwidth of 64-128kbps. MySQL [8] database system was used to store the set of substitution characters which were selected to match the original characters on basis of most common similar characteristics. The algorithm was deployed on a remote web server that used third party hosting services. An android phone was used as a client.

3.1 API Connectivity Codes

The sample code below demonstrates how to connect and interact with the web services API. The xmlseclibs project provides a lightweight PHP library for XML security. Prerequisite:

•   xmlseclibs (available at http://code.google.com/p/xmlseclibs/)

•   The exacttarget_soap_client.php file

```php
/* Set include directory for exacttarget PHP SOAP library files.
   The path to these files may already be set for the include_path
   setting within the php.ini file */
$includeDir = realpath(dirname(__FILE__).'/../includes/');
$path = ini_get('include_path');
$path = $includeDir . PATH_SEPARATOR . $path;
ini_set('include_path', $path);

require('exacttarget_soap_client.php');

/* Set location of WSDL */
$wsdl = 'https://.wsdl';

try {
    /* Create the Soap Client */
    @$client = new ExactTargetSoapClient($wsdl, array('trace'=>1));

    /* Set username and password here */
    $client->username = 'username';
    $client->password = 'password';

    /* Other Code Here */

}
```

The following code sample verified the connection to the ExactTarget API:

```php
$param = new ExactTarget_VersionInfoRequestMsg();
$param->IncludeVersionHistory = True;
$results = $client->VersionInfo($param);
var_dump($results);
```

4. RESULTS

There was successful implementation of the android app and the results tested, below are the output screens
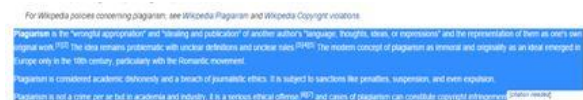
Step 1: Copying of contents

Fig 3: Copying a paragraph from a source
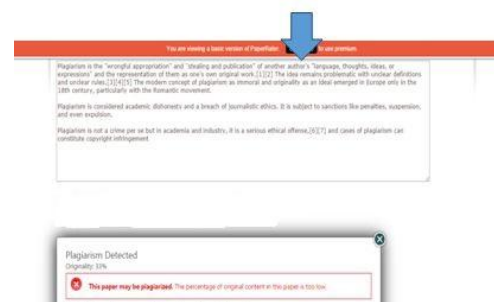
Step 2: Initializing the plagiarism check

Fig 4: Plagiarized Content (30% original)

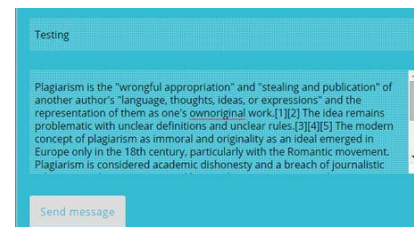Step 3: Entering text for processing

Fig 5: Submission for modification (Android App Interface)

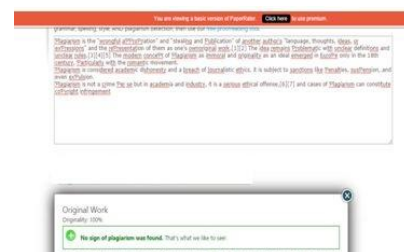Step 4: Modified unique contents

Fig 6: Results of modification (100% original)

It has also been tested for different types of text size and documented results in terms of time required for receive, processing   and   submission

.**Table 1. Content Results**

Text Size	Receive Time	Processing Time	Notification Time
50	0.5sec	0.7sec	1.2 sec
100	0.8sec	0.9sec	1.8 sec
200	0.9sec	1.5sec	2.1 sec

5. FUTURE WORK

This android application makes the user send the required text in a unique format. The overall aim is not to support the plagiarism or support/develop the methods that bypass plagiarism checks, but discover such possible ideas/techniques that will help the prevention of plagiarism bypass and avoid the detection.

In some of the results by following this approach of substitution the characters sometimes look slightly different or deformed from the English Language. Thus disguising the text or making new very similar fonts would make the changed characters more indescribable from the rest of the text.

6. CONCLUSION

According to certain statistics, over 2 billion people in the world speak English. It is the official language over 50 countries. Thus most universities and colleges all over the world use English as their medium of language. Also most of the electronics such as laptops etc., use an English alphabet keyboard. Thus most searches etc. are done in English. Many people are unaware of the fact that alphabets of different languages are present in the ASCII table. Thus there are alphabets in the English language that have similar looking alphabets in a different language with a different ASCII value. Thus can exploit this fact to the advantage and implement this in the application.

This also highlights the major flaw in majority of the plagiarism checkers available today, so this method can be utilized to rectify the flaw.

7. REFERENCES

[1] Brian Martin ,"Plagiarism: a misplaced emphasis", Journal of Information Ethics, Vol. 3, No. 2, Fall 1994, pp. 36-47

[2] Patrick M. Scanlon, David R. Neumann, "Internet Plagiarism Among College Students", Journal of College Student Development, VOL. 43 NO.3, JUNE 2002.

[3] "HTML 5", http://www.w3.org/TR/html5/, November,25, 2013.

[4] "PHP", http://php.net/, November, 28, 2013.Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[5] "Eclipse", https://www.eclipse.org/downloads/, December, 2, 2013.

[6] "Wikipedia", http://en.wikipedia.org/wiki/Main_Page, October, 12,2013.

[7] "Android", http://www.android.com/, October 25,2013.

[8] "MySQL", http://www.mysql.com/, September 26,2013.

8. Title and Authors

Mr. Pravin Shrinath is an associate professor in the Computer Engineering Department of the Mukesh Patel School of Technology Management & Engineering, NMIMS University, India.

Mr. Saurabh Malgaonkar is an assistant professor in the Computer Engineering Department of the Mukesh Patel School of Technology Management & Engineering, NMIMS University, India. Area of Interests: Networking, Cloud Computing. Email: saurabhmalgaonkar@gmail.com

Yash Radia is a 3rd year Computer Engineering student pursuing B-Tech Degree from Mukesh Patel School of Technology Management & Engineering, NMIMS University, India. Interested Domains: Web Development, Artificial Intelligence, Image Processing and Android. Email: yashradia.nmims@gmail.com

Nipun Philip is a 3rd year Computer Engineering student pursuing B-Tech Degree from Mukesh Patel School of Technology Management & Engineering, NMIMS University, India. Interested Domains: Data Warehousing, Android and Web Development. Email: pnipun125@gmail.com.