

Анализ современных систем обхода заимствований текста

Ересь Артём Владимирович

Приамурский государственный университет им. Шолом-Алейхема

студент

Лучанинов Дмитрий Васильевич

Приамурский государственный университет им. Шолом-Алейхема

старший преподаватель кафедры информационных систем, математики и методик обучения

Аннотация

В данной работе рассмотрены системы по обходу антиплагиата. Будут показаны способы обнаружения обработки этими системами и возможность вернуть исходный до обработки текст.

Ключевые слова: Антиплагиат, рерайтинг, работа с текстом.

Analysis of modern systems of circumvention of borrowings of the text

Yeres Artem Vladimirovich

Sholom-Aleichem Priamursky State University

student

Luchaninov Dmitry Vasilievich

Sholom-Aleichem Priamursky State University

Senior Lecturer

Abstract

In this paper, the system to bypass the anti-plagiarism. Will be shown ways to detect the treatment of these systems and the ability to revert to the default before processing the text.

Keywords: Anti-plagiarism, rewriting, working with text.

В настоящее время развитие информационных технологий привело к появлению большого объема информации, доступ к которой имеют все пользователи интернета. Появились сайты, накопившие в своей коллекции множество различных рефератов, лабораторных, курсовых и дипломных работ по различным темам. Это дало возможность пользоваться опубликованной информацией всем посетителям данных источников, при этом без изменения текста и указания первоисточника. Нередки примеры использования чужого материала без ссылок на первоисточники, что в научном сообществе считается недопустимым и влечет за собой ответственность.

В сфере образования преподаватели давно применяют в своей работе такой показатель как уникальность, позволяющий задать учащимся процент отличия текста работы учащегося от скопированного им текста из других источников. Для проверки уникальности текста придумана система под названием "Антиплагиат", позволяющая осуществить поиск заимствований в тексте из всех интернет-источников, заложенных в систему. Под плагиатом понимается умышленное присвоение авторства и заимствование текста чужой работы.

Различают несколько форм заимствований, встречающиеся в практике учебных заведений:

1. Полное или частичное копирование текста из одного источника.
2. Копирование и переработка текста из двух и более источников.
3. Копирование текста с использованием изменения порядка построения текста.

Разумеется прогресс не стоит на месте и появились методы обмана системы поиска заимствований для повышения уникальности текста. Это позволяет скрыть скопированный текст и выдать его за авторский. В данной работе мы проведем анализ современных систем обхода заимствований текста, рассмотрим способы обнаружения обработки этими системами и возможность возврата исходного до обработки текста

Целью данного исследования является анализ современных систем обхода заимствований текста.

Перед тем как приступить к анализу систем обхода заимствований, рассмотрим работы других авторов по проверке уникальности текста. В статье А.О. Зиберта и В.В. Мирошниченко описан процесс разработки алгоритма определения наличия заимствований в тексте с использованием эталонного множества слов. Рассмотрены достоинства и недостатки данного алгоритма, а также приведены различные методы создания эталонного множества слов [1]. Группа исследователей К.Ф. Сафин, М.П. Кузнецов и М.В. Кузнецова рассказали об определении заимствований в работах, где исследуемый текст анализируется изолированно, т. е. возможные источники заимствований неизвестны [2]. В своей статье И.К. Скокова представила проект, на внедрение системы проверки текстов на плагиат в образовательные учреждения [3]. Автор Е.А. Кучмурукова на примере наиболее известных интернет-сервисов характеризует принцип действия систем поиска заимствований, рассматривает их функциональные возможности, основные достоинства и недостатки [4]. Рассмотрела ререйтинг как новый инструмент созданий уникальных произведений и при этом проанализировала права автора О. Оськина [5]. Е.Н. Москаленко и Ю.Н. Слесарев рассмотрели способы реализации поиска плагиата с помощью сети интернет и внедрение этих способов для нового программного обеспечения [6]. В статье электронного источника рассмотрены способы обхода системы антиплагиат[7].

В настоящее время вопросы обхода антиплагиата и повышение уникальности являются достаточно популярными в сети интернет. Эта тема

интересует многих людей, обучающихся в различных учебных заведениях. Начнем с того, что за последние годы количество сервисов, позволяющих проверить текст на уникальность заметно выросло. Несмотря на то, что алгоритмы поиска заимствований у этих программ разные, все равно результат проверки уникальности отличается на несущественные доли процента.

Начнем с того, что подробно разберем способы обхода программы Антиплагиат. Их можно разделить на технические способы повышения уникальности и рерайтинг. Первая группа связана непосредственно с изменением текста, всё делается максимально быстро, с расчетом на обман системы, но и риск разоблачения более высок. Существует множество технических методов повышения уникальности, но из-за совершенствования программ поиска заимствований некоторые уже не работают и не используются.

Рассмотрим некоторые из них:

1. Замена русских букв в тексте на иноязычные (в основном английские). Метод не работает так как, буквы для системы воспринимаются по внешнему виду, не обращая внимание на кодировку. Многие современные сервисы поиска заимствований при выдаче результата покажут, что в тексте имеются замены символов, указав при этом точное их количество. К тому же при сдаче электронного варианта работы преподавателю, в тексте Word замены символов будут подчеркиваться.

2. Замена пробелов символами, буквами и точками, после чего закрашивание этих мест белой заливкой. Этот метод не работает, потому что системы антиплагиата сообщают о подозрительности документа, что заставит проверяющего более тщательно изучить отчет.

3. Автозамены с использованием скрытого переноса либо другого символа. С недавнего времени системы поиска заимствований обращают внимание и на этот метод, поэтому уникальность не повышается.

4. Перестановка слов и предложений местами, разбиение сложных предложение на несколько и наоборот. Способ совершенно не влияет на уникальность.

Как видим, совершенствование системы поиска заимствований происходит быстро и некоторые методы уже не актуальны. Но все же есть способы еще работающие в настоящее время. Они связаны с некоторыми системами работающими против систем антиплагиата, так называемые антиплагиат-киллеры. Рассмотрим способы которые они используют.

1) Смена кодировки текста. Это позволяет сделать текст уникальным.

2) Замена некоторые слов на формулы, которые антиплагиат не воспринимает вообще.

3) Вставка в файл Word другие незаметные внешние символы, которые отчет антиплагиата не замечает, но уникальность повышается вполне реально.

Так же для повышения уникальности можно использовать самые простые методы такие как замена таблиц и некоторых символов картинками;

использование встроенной функции Word, позволяющей вставить картинку, график или диаграмму; замена текста функцией. Эти способы могут помочь повысить уникальность вашего текста.

Все же использование некоторых методов не дает полной гарантии результата, банальная проверка на нескольких системах антиплагиата может вычислить обман. В то же время использование антиплагиат-киллеров имеет большие риски. В сети интернет масса программ, именуемых таким названием, но при скачивании это может оказаться совсем не то, что вы искали. Установка нежелательного программного обеспечения может повредить вашему компьютеру. Приобретенные ранее версии антиплагиат киллеров со временем могут перестать работать, так как системы поиска заимствований не стоят на месте и развиваются. Так же существует много предложений в сети интернет о программах повышения уникальности, на непроверенном источнике можно легко нарваться на мошенников и понести материальный убыток.

Для повышения уникальности можно использовать более честные способы, такие как рерайтинг, копирайтинг и перевод с иностранных источников.

Рерайтинг – это переработка текста своими словами из одного или нескольких источников. К примеру, используя начало из одного источника, план повествования из другого и заключение из третьего, при этом перефразировав текст своими словами, можно получить уникальный текст. При глубоком рерайте, когда человек действительно разбирается в той сфере к которой относится текст, можно получить абсолютно уникальную работу.

Копирайтинг является более сложным способом, требующим от человека более серьезного понимания текста и хороших умений в пересказе. Для этого метода необходимо прочитать текст из какого-либо источника, хорошо понять его, и пересказать своими словами. Копирайтинг – это более качественно выполненная работа. При проверке на заимствования результат будет более высок, чем при рерайте. Текст имеет большую уникальность, так как написал с нуля, используя смысл из источника.

Неплохим способом для обмана системы поиска заимствований является использование иностранных источников. Антиплагиат не сможет проверить текст на перевод, а вам нужно лишь найти иностранный источник, перевести и привести текст в более осмысленный вид. Этот способ позволяет достаточно хорошо повысить уникальность вашей работы.

Вопросы обнаружения обработки методами повышения уникальности были подняты в нашей работе ранее. Рассмотрим вопрос возврата исходного текста, который был до обработки. При использовании рерайтинга и копирайтинга человек заинтересованный в поиске исходного текста, может использовать отрывки из текста для поиска в сети интернет. Есть вероятность того, что по отрывкам и смыслу будет найден исходный до обработки текст. Для проверки на перевод можно перевести текст работы и поискать в сети интернет, используя предложения из него. Возможно, текст будет найден и обман будет разоблачен. Для возврата исходного текста при использовании

технических методов можно использовать функцию автозамены и попробовать замену различных символов в Word. В некоторых случаях это может помочь. В основных случаях нужно использовать поиск в сети интернет по тематике, отрывкам из текста и предложениям.

Таким образом, мы рассмотрели возможные способы по обходу системы поиска заимствований в тексте. Показаны способы обнаружения обработки этими системами и возможность вернуть исходный до обработки текст. В результате анализа можно сказать, что такие системы существуют, но в действительности некоторые из них не приносят никакого результата, а лишь некоторые из них реально помогают повысить уникальность текста.

Библиографический список

1. Зиберт А.О., Мирошниченко В.В. Современные методы определения наличия заимствований в тексте // Научные исследования: от теории к практике. 2015. Т. 2. № 2 (3). С. 41-42.
2. Сафин. К.Ф., Кузнецов М.П., Кузнецова М.В. Определение заимствований в тексте без указания источника // Информатика и ее применения. 2017. Т. 11. № 3. С. 73-79.
3. Скокова И.К. Разработка проекта на внедрение АИС для проверки текстов на заимствования // Экономика и социум. 2014. № 3-4 (12). С. 98-106.
4. Кучмурукова Е.А. Системы антиплагиата как инструмент для проверки оригинальности текста // Библиография и книговедение. 2015. № 3 (3). С. 17-21.
5. Оськина. О. Ререйтинг как новый инструмент создания произведения. Защита прав автора при ререйтинге // Интеллектуальная собственность. Авторское права и смежные права. 2014. № 10. С. 43-48.
6. Москаленко Е.Н., Слесарев Ю.Н. Методы поиска заимствований в сети Интернет // Современные научные исследования и инновации. 2017. № 4 [Электронный ресурс]. URL: <http://web.snauka.ru/issues/2017/04/80894> (дата обращения: 12.01.2018).
7. Система antiplagiat.ru и способы её обхода URL: <http://mastersloga.ru/sistema-antiplagiat/> (дата обращения 12.01.2018)
8. Кизянов А.О., Ленкин А.В., Лучанинов Д.В. Реализация локального антиплагиата с помощью объектного приложения языка python // Постулат. 2016. № 1 (1). С. 17.