**АКАДЕМИЯ НАУК РЕСПУБЛИКИ ТАДЖИКИСТАН**

**ИНСТИТУТ МАТЕМАТИКИ ИМЕНИ А. ДЖУРАЕВА**

**РЕФЕРАТ**

**Тема:** **«***Проверка математических формул в* Latex*-е на наличие плагиата***»**

PhD 1-го курса института математики АНРТ

Направление «информатика»

Мирзоев Абубакр Шарафович

2019 год

**СОДЕРЖАНИЕ**

Введение

Что такое плагиат

Методы обнаружения плагиата

Математические символы, формулы в Latex-е

Алгоритм выявления плагиата между формулами

Преобразование формулы в короткую запись

Сравнение коротких записей

Реализация алгоритма в ЯП JavaScript

Заключение

Список литературы

**ВВЕДЕНИЕ**

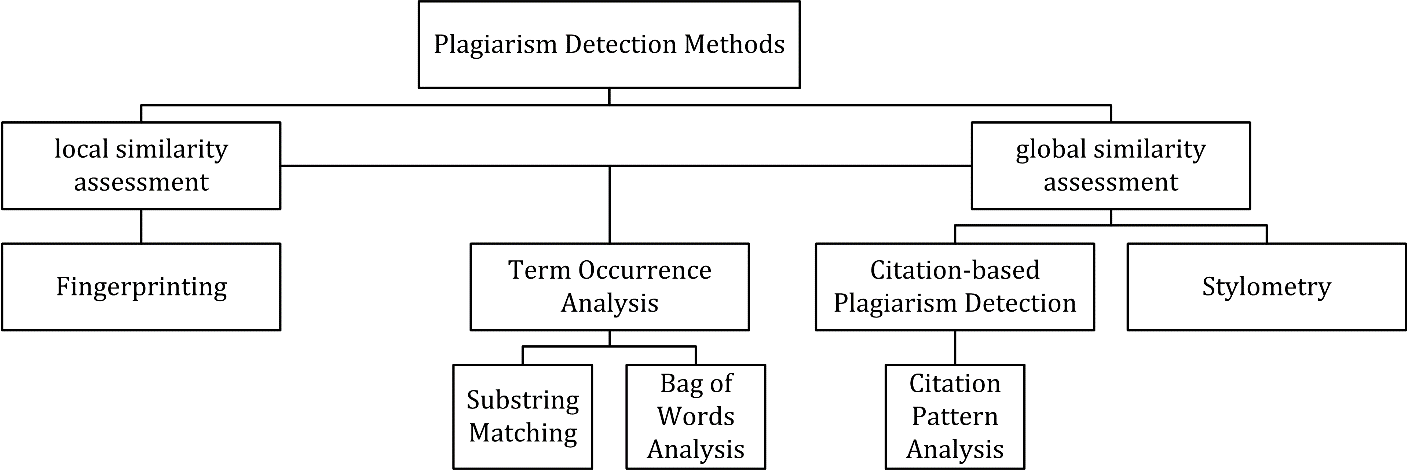
Целью данной работы является получение ответа на вопрос наличия плагиата в математических работах, а именно, в формулах.

Чтобы создать алгоритм решения такой задачи необходимо понять какие методы выявления плагиата существуют и в чем их преимущества и недостатки и, который из них больше подходит для нашей задачи.

**ЧТО ТАКОЕ ПЛАГИАТ**

По определению из Википедии[[1]](#footnote-1) **плагиат** – это использование, перефразирование и подведение итогов работы в любой форме без подтверждения ссылками на источники и представление ее как своей собственной работы.

**Методы обнаружения плагиата**



**Соответствие подстроки**

Шаблон поиска “needle”, строка, в которой ведется поиск “haystack”. Кроме того, дан алфавит, на котором проводится поиск. На данный момент придуманы много алгоритмов поиска подстроки в тексте.

**Анализ «множества слов»**

Принцип метода:

Строка представляется «множеством ее слов», без учета *грамматики* и *порядка* слов. Этот метод в основном используется для классификации документов, где *частота* появления слова используется как признак для тренировки *классификатора*.

(**Цитирование и Стилометрию** можно посмотреть в [википедии](htttps://wikipedia.ru))

**МАТЕМАТИЧЕСКИЕ СИМВОЛЫ И ВЫРАЖЕНИЯ В LATEX-Е**

Подробнее с командами латех можете ознакомится в интернете, многие из них записаны в коде **js\_realization**

**АЛГОРИТМ ВЫЯВЛЕНИЯ ПЛАГИАТА МЕЖДУ ФОРМУЛАМИ**

**Преобразование формулы в короткую запись**

**Цель:** из длинного выражения сделать короткую для облегчения сравнения.

Пример, из выражения «**\frac12+\frac x 2 = \frac{1+x}2**» (предположим номер элемента массива со значением **frac** равен 1. Заменить надо на **el1**, то есть el+номер-элемента-в-массиве)

сделать массив «**el1, 12, +, el1, x, 2, =, el1, {, 1, +, x, }, 2**».

**Алгоритм:**

1. Все математические **символы** и **операции** (, которые пишутся после знака **«\»** (обратный слэш)) нужно ***записать*** в массив, ***отсортировать*** по *алфавиту* и *количеству* символов; [done]
2. Входное выражение записать в массив посимвольно;[done]
3. [done]
   1. Создать новый массив для короткой формы;
   2. Проверять символы входного массива, если символ равенодному из простых символов (+, -, \*, /, скобки), то записать его в новый массив, иначе, если равен «\» то выяснить что идет после этого знака, если это один из видов скобок, то пропустить, иначе найти где заканчивается (признаки конца - скобки, \_, \) название операции или символа. Вырезать название и искать его в массиве символов и операций;
   3. Найти номер этого символа или операции по названию (первая буква) и длине (количество символов) и записать его в массив как **el**+**номер**

**Сравнение коротких записей**

Сравнение будет поэлементным, то есть номера сравниваемых элементов одинаковы.

Процент плагиата вычисляется по следующей формуле:

Назовем сравниваемые выражения ВырА и ВырБ, тогда

ПроцПлаг = (КолОдинЭл) / Макс(Длина(ВырА), Длина(ВырБ))

**РЕАЛИЗАЦИЯ АЛГОРИТМА В ЯЗЫКЕ ПРОГРАММИРОВАНИЯ JAVASCRIPT**

1. <https://ru.m.wikipedia.org/wiki/Выявление_плагиата> [↑](#footnote-ref-1)