

ПРИНЦИПИ РОБОТИ ПРОГРАМ ПЕРЕВІРКИ НА ПЛАГІАТ

В даний час існує достатня кількість програм, що визначають рівень оригінальності тексту робіт. Програми перевірки на плагіат не показують, списана ця робота з готових робіт з інтернету чи ні. Програми показують тільки запозичений текст. Запозиченим текстом програми вважають будь-який текст, знайдений в інтернеті або у власних базах.

Програми для виявлення плагіату можуть використовувати шинглові, кореляційні алгоритми або обидва відразу. Алгоритм шинглів дозволяє швидко знаходити запозичення в дуже великих базах документів. Цей алгоритм заснований на розбитті тексту документа на невеликі послідовності слів певної довжини – шингли. Зазвичай використовується шингли довжиною від 4 до 6 слів. Для кожного шинглу розраховується значення хеш-функції. Пошуковий індекс формується як відсортований список значень хеш-функції з вказівкою ідентифікаторів документів, в яких зустрілися відповідні шингли. Документ, що перевіряється, також розбивається на шингли. Потім за індексом знаходяться документи з найбільшою кількістю збігів по шинглах з документом, що перевіряється.

Такий алгоритм дозволяє швидко виявляти запозичені фрагменти. Однак труднощі виникають для фрагментів тексту в випадках запозичення з іноземних джерел після машинного перекладу. Як закордонні джерела в основному використовують англійські статті, тому що англійською мовою написана досить велика кількість текстів, вчені в більшості випадків як другу «робочу» мову використовують саме англійську, більшість міжнародних наукових журналів видається англійською мовою.

Для подолання цих труднощів підключаються кореляційні алгоритми, які дають можливість перевіряти тексти на схожість за змістом.

Пошук «за змістом» забезпечується через кластеризацію англійських слів так, щоб семантично близькі слова та словоформи одного і того ж слова потрапили в один кластер. Після цього, перед розбиттям текстів на шингли, слова замінюються на мітки кластерів, до яких ці слова належать.

Важливо відзначити, що програми знаходять джерело запозичення, але не визначають, чи є воно першоджерелом.

О. В. Чаленко, О. В. Казанко

ОБ'ЄКТИВНІСТЬ ОЦІНЮВАННЯ ЗНАНЬ. ПСИХОЛОГІЧНІ ТА МЕТОДИЧНІ АСПЕКТИ

Успішність кожного студента повинна бути оцінена з максимальною точністю й об'єктивністю. Оцінка в очах студента мусить бути справедливою і переконливою. З цією метою необхідно вилучити випадки