Секція 1

АНАЛІЗ МЕТОДІВ ВИЯВЛЕННЯ АВТОРІВ ФАЙЛУ, ЩО ДОСЛІДЖУЄТЬСЯ

Малєєва З.-Т. О
Національний аерокосмічний університет ім. М. Є. Жуковського «ХАІ»
Науковий керівник: Пєвнєв В. Я.

Актуальність. З розвитком мережі Інтернет цифрові дані стали невід'ємною частиною повсякденного життя, проблема ідентифікації авторства файлів набула важливого значення в різних галузях, включно з комп'ютерною криміналістикою, кримінальним розслідуванням та авторським правом.

Ця проблема вимагає розробки ефективних методів ідентифікації авторства файлу, які допоможуть визначити, хто несе відповідальність за створення або редагування конкретного файлу. Аналіз авторства - це статистичне дослідження лінгвістичних та обчислювальних характеристик письмових документів окремих осіб [1].

Відомі дослідження аналізу авторства для текстової комунікації в Інтернеті є дуже обмеженим. Традиційні письмові твори є об'ємними і, як правило, добре структуровані відповідно до загальних синтаксичних і граматичних правил. На відміну від них, онлайн-документи, такі як електронні листи та миттєві повідомлення, короткі, погано структуровані і зазвичай написані мовою абзаців, що містить кілька орфографічних і граматичних помилок. Ці відмінності роблять деякі традиційні роботи з аналізу авторства неможливими для застосування до текстових даних в Інтернеті.

Метою даної роботи є дослідження методів виявлення авторства файлів, отриманих в мережі Інтернет.

Основні положення. У доповіді надано дослідження авторства, яке показують, що окремі особи часто залишають сліди їх особистості у своїх письмових роботах [2]. Наприклад, вибір слів, композиція речень і абзаців, а також відносна перевага одних мовних артефактів над іншими можуть допомогти відрізнити одну особу від іншої. Перехопивши та проаналізувавши «почерк» анонімних текстових повідомлень, є можливість виокремити певні характеристики одного автора повідомлень з вибірки.

При ідентифікації авторів файлів отриманих в інтернеті, використовують різні підходи та методи, а саме: стилістичний аналіз тексту, виявлення місця розташування відправника, експертна оцінка, порівняння зразків авторства, машинне навчання.

Наведені у доповіді методи визначення автора файлів показує, що це завданням є непростим і багато авторів намагаються приховати свою

особистість, змінюючи стиль [3]. Крім того, неможливо з упевненістю визначити автора файлу, аналізуючи тільки текст і метадані, оскільки інші автори можуть мати схожий стиль або доступ до аналогічної інформації. Але використовуючи поєднання деяких з цих методів може дати позитивний результат у визначенні автора конкретного файлу.

Висновки. Представлені у доповіді методи та їх комплексне використання можуть допомогти в ідентифікації автора. Однак, як на сьогодні не існує консенсусу щодо правильної методології. Майже кожен з представлених методів страждає від таких проблем, як сумнівний аналіз, непослідовність для одного і того ж набору авторів, невдала реплікація. У таких випадках головну роль починає відігравати кваліфікація експерта. Користувачі повинні дотримуватись організаційних методів безпеки та не відкривати файли невідомого походження.

Список літератури

1. R. H. Baayen, H. van Halteren, and F. J. Tweedie. Outside the cave of shadows: using syntactic annotation to enhance authorship attribution. URL - http://www.sfs.unituebingen.de/hbaayen/publications/BaayenHalterenTweedie1996.pdf (дата звернення: 12.11.2023);
2. O. de Vel, A. Anderson, M. Corney, and G. Mohay. Multi-topic e-mail authorship attribution forensics. URL – https://eprints.qut.edu.au/8039/1/8039.pdf (дата звернення: 13.11.2023);
3. A Unified Data Mining Solution for Authorship Analysis in Anonymous Textual Communications. URL – https://spectrum.library.concordia.ca/id/eprint/976945/1/fung2011b.pdf (дата звернення: 13.11.2023);

Відомості про авторів

Малєєва Злата-Тіна Олександрівна, магістрантка кафедри комп'ютерних систем, мереж і кібербезпеки, НАУ «ХАІ», z.malieieva@student.csn.khai.edu
Пєвнєв Володимир Яковлевич, професор кафедри комп'ютерних систем, мереж і кібербезпеки, НАУ «ХАІ», д.т.н., доцент, v.pevnev@csn.khai.edu