

ФОРМАЛЬНІ МЕТОДИ АВТОРОВІДЧЕСЬКОЇ ЕКСПЕРТИЗИ

За останні роки різко зросло число різних конфліктних ситуацій, в яких інструментом і засобом правопорушення є слово звучне, надруковане або написане, відображене на будь-якому носіїв і поширюване будь-яким засобом, включаючи масову комунікацію в Інтернеті.

Порушення авторських і суміжних прав, прав на фірмові найменування і товарні знаки, доменні імена, авторські матеріали, що передаються і розміщені на інформаційних ресурсах Інтернету, викликало потребу в проведенні нових пологів експертиз, що відносяться до класу судово-мововідческих. Це – авторовідческа лінгвістична і фоноскопична експертиза.[1, 2]

Дана робота присвячена розгляду питань авторовідчеської експертизи.

Авторовідческа експертиза - один з різновидів ідентифікації особи, в процесі проведення якої здійснюється ідентифікація (встановлення) автора спірного тексту, встановлення його статі, віку, професії, рівня освіти і ряд інших фактів, встановлення приналежності листування конкретним історичним особам і ін.

Формальні методи авторовідчеської експертизи засновані на порівнянні обчислюваних характеристик текстів. У загальному випадку текст відображається у вектор обчислених для нього параметрів, кожен з яких об'єктивно характеризує деякий набір особливостей тексту. Таким чином, текст графічно відображається в деяку точку n-мерного простору. При такій формалізації автор також може бути представлений у вигляді аналогічного вектора параметрів – цим вектором буде вектор текстів, написаних даним автором.

В більшості випадків як характеризуючі параметри тексту вибираються ті або інші його статистичні характеристики: статистика використання певних частин мови, деяких конкретних слів, розділових знаків, фразеологізмів, архаїзмів, рідкісних і іноземних слів, довжина речення (у словах, складах, знаках) і так далі.

Для підвищення якості проведення авторовідчеської експертизи можна запропонувати наступну методіку [3, 4]. Пошук числа повторень тієї або іншої службової частинки серед тисячі слів досліджуваного тексту ототожнюваний з відомим завданням математичної статистики про повторення випробувань, тобто кількість слів тексту вважатимемо числом випробувань, а кількість m_i повторень частинки – числом появ події. Тоді можна ввести поняття частоти як відношення вказаних чисел:

$$P_i = \frac{m_i}{n_i} \quad (1)$$

У математичній статистиці відомі випадки, коли при збільшенні числа випробувань числові значення частот коливаються біля певної величини і відхилення частот від вказаної величини зменшуються із зростанням числа випробувань. Як правило, такою величиною приймається середнє арифметичне частот P_i . Якщо у формулі (1) символом i позначатимемо номер серії випробувань, то P_{cp} необхідно обчислювати так:

$$P_{cp} = \frac{\sum_i P_i}{N} \quad (2)$$

де N — число серій.

У статистиці описаний факт повторюваності частот називається законом стійкості частот, а на основі відомої теореми Я. Бернуллі величина P_{cp} приймається як вірогідність появи розшукуваної події.

Для виявлення закону стійкості частот стосовно головного приємникового спектру (y, na, z) весь даний текст розділяється на фрагменти з тисячі слів. Далі починаємо розшукувати кількість повторень кожного приємника даного спектру в першій тисячі слів. В цьому випадку число випробувань $n_1 = 1000$ назвемо першою серією випробувань, а отримане число m_1 повторень даної службової частинки слід вважати числом появ розшукуваної події. Тепер по формулі (1) можна обчислити частоту P_1 першої серії випробувань. Для отримання частоти P_2 другої серії необхідно до першого фрагмента тексту додати другий і для $n_2 = 2000$ з урахуванням нового значення m_2 обчислити P_2 за формулою (1). Вказаний процес продовжити до тих пір, поки вказаним аналізом не буде охоплений весь досліджуваний текст. Отриманий таким чином набір чисел P_i покаже, чи має місце закон стійкості частот. Якщо закон стійкості частот матиме місце, то середню частоту, що характеризує вірогідність появи даної службової частинки, обчислимо за формулою (2).

Література:

Баранов, А.Н. Лингвистическая экспертиза текста : теория и практика : учеб. пособие / А.Н. Баранов. — М. : Флинта : Наука, 2007. — 592 с.

Бринев, К.И. Теоретическая лингвистика и судебная лингвистическая экспертиза / К.И. Бринев. — Барнаул, 2009.

О.В. Кукушкина, А.А. Поликарпов, Д.В. Хмелев Определение авторства текста с использованием буквенной и грамматической информации / Проблемы передачи информации. — Т. 37. — № 2. — 2001.

В.А. Бубнов, А.В. Сурвило Формальный анализ авторской индивидуальности литературного произведения / Информатика и образование — М. — 2007. — № 8. — С. 83-87.