

УДК 343.148/343.533

Г. С. Римарчук

АЛГОРИТМИ ДОСЛІДЖЕННЯ ВЕРБАЛЬНОЇ СХОЖОСТІ ЗНАКІВ ДЛЯ ТОВАРІВ ТА ПОСЛУГ

Постановка проблеми. Нерідко для встановлення схожості чи відрізняльної здатності торгової марки призначають судову експертизу. Експерт, так само як і юрист чи працівник патентного відомства, зобов'язаний провести детальну та ретельну експертизу на наявність схожих до ступеню змішування знаків. При цьому необхідно вибрати правильну комбінацію алгоритмів пошуку для його ефективності та точності результату.

Мета публікації. Дослідження вербальної схожості знаків для товарів та послуг повинно включати в себе системний підхід до пошуку аналогів чи подібностей позначень, тобто мається на увазі поетапне дослідження самого знаку, вивчення існуючих баз даних (або використання однієї профільної) та встановлення необхідних критеріїв пошуку.

Стан дослідження. Питанням дослідження вербальної схожості знаків для товарів та послуг займалися такі вітчизняні та зарубіжні автори-науковці: І. Кириченко, О. Колосов, Б. Бродда, В. Бугдал, П. Дарт, Д. Холмс, Л. Филипс, Дж. Зобель та ін.

Виклад основного матеріалу. Надання правової охорони знаку для товарів/послуг забезпечує першочергове право власника свідоцтва на захист створеного об'єкту індивідуалізації виробництва, а саме словесного позначення, логотипу, зображення чи комбінації словесного та зображувального елементів, що має відрізняльну здатність серед інших позначень на ринку товарів і послуг, від неправомірних дій із боку третіх осіб. Окрім традиційних об'єктів, які є носіями словесного чи зображувального змісту, правовій охороні підлягають також колір позначення, 3-вимірні позначення, звукові марки та позначення, які сприймаються на запах.

Передусім торгові марки повинні набути відрізняльної здатності, а саме стати беззаперечним індикатором походження товару/послуги. У протилежному випадку існує явна загроза щодо визначення джерела походження товарів/послуг та можливість виникнення конфлікту позначень. Закон України «Про право на знаки для товарів та послуг» [1, с. 36] містить вказівку на можливість відмови в реєстрації торгового позначення на підставі визнання такого знаку ідентичним позначенню, яке отримало правову охорону раніше, тобто за датою пріоритету цей знак є більш раннім, а також відповідає аналогічній категорії товарів/послуг, для яких призначена торгова марка. Окрім цього, береться до уваги також «відносна ідентичність» – подібність із раніше зареєстрованою торговою маркою та категорією товарів та послуг, що у свою чергу створює ймовірність сплутання торгових знаків, оскільки ступінь їх змішування є достатньо високим.

З метою встановлення розрізняльної здатності торгових знаків необхідно провести низку комплексних досліджень, при цьому детально вивчаючи базу даних існуючих територіальних торгових марок. Функцію пошукової системи насамперед виконують національні патентні відомства, вони ж інформують заявників про зареєстровані та такі, що подані на реєстрацію, торгові марки, а також здійснюють моніторинг нових

заявок для іноземних заявників. Державна служба інтелектуальної власності України надає доступ до баз даних зареєстрованих торгових марок та добре відомих знаків в Україні. Такими пошуковими інструментами є бази даних «Зареєстровані в Україні знаки для товарів та послуг» та «Відомості про добре відомі знаки в Україні» [2].

Виконання пошуку на схожі знаки для товарів та послуг повинно керуватися категоріями товарів/послуг, для яких призначатиметься торгова марка. Із цією метою використовується розроблена Всесвітньою організацією інтелектуальної власності Ніццька класифікація товарів та послуг. Ця класифікація значно полегшує та звужує простір для пошуку до необхідного числа торгових марок.

Найбільш розповсюдженими є знаки, що складаються зі словесних позначень. Такі позначення визнаються схожими, якщо фонетична (звукова), візуальна (графічна) чи семантична (смилова) схожості, встановлені на підставі аналізу відрізняльних ознак, створюють передумову до того, що споживач, знаючи один знак, буде сприймати його як інше позначення (тобто виникнення асоціації з іншим позначенням). Зазвичай в таких випадках призначається судова експертиза, у ході якої вивчається семантичне навантаження слів, особливості їх звучання та графічного виконання [3, с. 82].

Відтак схожість торгових марок повинна визначатися за вербальними та зображувальними ознаками. Зображувальна схожість торгових марок встановлюється шляхом пошуку та визначення коду відповідного торгового знаку Віденської класифікації – ще одного інструменту класифікації товарів та послуг ВОІВ [4]. Віденська класифікація являє собою ієрархічну систему, що складається зі 114 категорій та 1 887 секцій, і, відповідно, всі існуючі торгові марки, які мають у своєму складі зображувальний елемент, віднесені до тих чи інших категорій з описом самого зображення. Результатом співпадіння коду Віденської класифікації нової торгової марки з кодом, який вже існує в базі даних, є перелік схожих торгових марок.

Словесна схожість визначається за допомогою порівняння слів чи словосполучень, заявлених на реєстрацію, з усіма існуючими словесними елементами торгових марок. Пошук словесної схожості в базах даних торгових марок повинен здійснюватися відповідним чином, беручи до уваги досить широкий спектр схожих марок. Питанням фонологічного аналізу торгових марок займався у свій час науковець Б. Бродда для Шведського патентного агентства (1960–1970 рр.) [5, с. 405–409]. Водночас він визначив підхід, за яким необхідно встановлювати фонологічні відмінності. Словесна схожість включає в себе фонетичні спорідненості та, крім цього, синоніми та перестановки букв у словах. При цьому параметри пошуку нагадують програму перевірки орфографії, а саме системи отримання фонетично схожих прізвищ у телефонних довідниках. Такі системи розроблені з метою виявлення слів та символів, які дещо змінені, зважаючи на послідовність результатів пошуку. Отримані слова можуть містити додаткові букви, переставлені чи видалені знаки.

У таблицях 1 і 2 ми наводимо приклади проведеного пошуку правового конфлікту, який мав місце в Швейцарії. Дані таблиці показують, у яких випадках було визнано схожість до ступеня змішання торгових марок, а в яких торгові марки були визнані самостійними об'єктами з можливістю отримання правової охорони. Загалом позначення вважатимуться словесно схожими з декількох причин: звучання, написання та смислова близькість [6, с. 135–143].

Таблиця 1

Схожість

FMH	FNH
PRETON	PREBETON
KISS	SOFT-KISS
7SEVEN	SEVEN PICTURES
RED BULL	BLUE BULL
YOUNET	Y YOUNET
CYREL	CYRA
MONA	MONO
FLY AWAY	FLOAT AWAY
DOGO	DO & CO

Таблиця 2

Відмінність

SEROPRAM	CITALOPRAM
MONTREUX	TСВ-THE MONTREUX
CELEBRY	BIC CELEBRATE
CRUNCH	SNICKERS CRUNCHER
ACTIVIA	ACTEVA
NUTRICIA	NUTRI
S.W.I.F.T.	SWIX
MARCO POLO	POLO

Відтак торгові марки зі схожим фонетичним наповненням можуть спричинити незручності для споживачів через труднощі в ідентифікації джерела походження товарів/послуг (як наприклад СТАРКС і СТАКС). Однак вимова того чи іншого позначення не завжди може бути очевидною чи унікальною, оскільки зустрічаються вигадані слова для позначення торгової марки (КІТАН) або слова іноземного походження.

Торгові марки, визнані схожими, часто є такими через використання певних елементів: чисел замість слів, додавання чи видалення букв або складів. Так, наприклад, схожими можна вважати марки, у позначенні яких має місце заміна однієї букви (ФОТОН і ФАТОН, 777 та Д77) або видалення букви (БЕЙЛІ та БЕЙЛ). Однак, з іншого боку, коригування однієї букви не завжди стане причиною схожості, особливо це стосується коротких позначень: SFS та TFS. Схожість також можна встановити, якщо позначення нараховують однакову кількість складів та водночас починаються чи закінчуються однаково: МОБІЛЕНД та МОБІТЕЛ, GRILON та GELON.

Ідейна схожість може стосуватися перекладу, синонімів чи антонімів: RED BULL і BLUE BULL). Торгові марки часто формуються з фрази чи слогану. Тому додаван-

ня одного слова може зіграти ключову роль у визначенні схожості торгових марок (наприклад, МЕДІ ЛАБ та МЕДЛАБ ШВЕЙЦАРІЯ, але VISA і JET-SET VISA).

Наведені вище приклади показують нам складність виведення єдиних правил встановлення схожості торгових позначень. Учений Б. Бугдал (B. Bugdahl) вивів 18 критеріїв визначення схожості словесних елементів торгових марок. Ці правила включають у себе однакові слова, слова, які починаються чи закінчуються однаково, слова з однаковими голосними та інші ознаки, які базуються на довжині слова та кількості букв, які співпали [7, с. 169–180].

Єдиний алгоритм пошуку словесної схожості позначень важко підібрати, однак існують загальні правила: вказавши в пошуковій системі слово-позначення торгової марки, система пошуку видає перелік схожих позначень. Можна виділити два основні види пошуку словесної схожості слів: визначення схожості позначень через встановлення порядку букв у слові та визначення схожості через фонетичні ознаки.

Спочатку розглянемо алгоритм Левенштейна – так звану редакторську відстань. Для оцінки схожості прикладів, що мають послідовну структуру, часто використовують відстань Левенштейна [8, с. 2], тобто вихідні позначення порівнюють з уже існуючими в базі на основі визначення одноелементних вставок, видалень та перестановок, необхідних для того, щоб результат пошуку внаслідок таких трансформацій відповідав початковому позначенню. Наприклад, ARVAL вважатиметься схожим з AREAL (відстань 1: перестановка однієї букви (елементу позначення)), MARTAL (відстань 2: одна вставка та одна перестановка) або ж TOPVAL (відстань 3: одна вставка та дві перестановки).

Інший алгоритм пошуку базується на використанні послідовності N-елементів, так званих N-грам. Суть цього пошуку полягає в співставленні послідовності букв, складів, слів чи звуків, що власне і є N-грамами [9, с. 91–130]. Наприклад, позначення ARVAL поділено в такій послідовності: ^A, AR, RV, VA, AL, L\$. У цьому ланцюжку символи ^ та \$ виступають заміниками початкових та кінцевих символів слова та служать показником визначення додатково можливих N-грам, на основі кількості яких здійснюється оцінка схожості позначень. Наприклад, позначення ARVAL вважатиметься схожим із позначенням PARVAL та ARAL, у яких буде відстань 3 [10, с. 331–345].

Недоліком цього алгоритму служить нездатність програми до фонетичного співставлення таких слів, як, наприклад, SYFA і СІРНА.

Для словесних позначень фонетична (звукова) схожість встановлюється на підставі таких ознак [3, с. 82–83]:

- наявність близьких чи співпадаючих звуків у позначеннях, що порівнюються;
- близькість звуків, що складають позначення, які порівнюються;
- розташування близьких звуків та звукосполучень відносно один одного;
- наявність співпадаючих складів та їхнє розташування;
- кількість складів у позначеннях;
- місця співпадаючих звукосполучень у позначенні;
- близькість складу голосних;
- близькість складу приголосних;
- характер співпадаючих частин позначень;

- входження одного позначення в інше;
- наголос.

Найдавнішим фонетичним методом пошуку можна вважати розроблений у 1918 році алгоритм «Саундекс» (Soundex) для індексації звуків із метою виявлення однакових за звучанням слів. При цьому пошукове слово та відповідна база даних зашифровані в коди саундекса. Такий код формується на основі першої букви слова та наступних (максимум трьох) цифр, які позначають букви продовження цього слова відповідно до таблиці кодів та букв [11, с. 166–172].

Таблиця 3

Коди саундекса

0	AEIOUYHW
1	BFPV
2	CGJKQSXZ
3	DT
4	L
5	MN
6	R

Таким чином, ARVAL буде зашифрований як A614, так само як і ARBELL, AIRPLUS, AIRBALL. Однак відмінність між словами доводиться перевіряти вручну через велику кількість схожих за кодуванням слів.

Наступний алгоритм призначений для кодування англійських слів з урахуванням фонетичного складу позначення, тим самим зменшивши їх до 12 звуків плюс «А» як індикатор першого голосного звуку [12, с. 38–43]. Наприклад, обидва позначення ARVAL і ORVILLE будуть зашифровані фонетично як ARFL. Окрім того, цей алгоритм дозволяє створювати альтернативне фонетичне шифрування для складених слів: MOBIGEL як MPJL та MPKL.

Нарешті, останнім у нашому дослідженні є алгоритм пошуку «Едітекс» (Editex). Цей алгоритм пошуку виділяє дев'ять фонетичних груп, які розподіляють усі приголосні звуки на кілька груп та всі голосні в одну окрему групу. Суть пошуку така ж, як і в алгоритму редакторської відстані, з поправкою на зазначене групування звуків. Алгоритм створений із метою встановлення найбільш фонетично подібних позначень та був протестований під час пошуку схожих прізвищ людей [13, с. 22–27]. Пошукові алгоритми, які націлені на отримання синонімічного ряду до заданого слова, вимагають застосування багатомовного тезауруса. Такий підхід зумовлює деякі незручності пошуку, оскільки результатом може стати або завелика, або замала кількість слів.

Пакети (бази даних) у вільному доступі та комерційні пакети пошуку, такі як dtSearch і Verity, дозволяють здійснювати пошук як відносної схожості (неточний пошук), так і словесного ланцюжка (різні форми слова, змінні закінчення) на фонетичному рівні (наприклад, CONNECT, CONNECTING, CONNECTED). Для використання таких баз даних необхідно прикласти чимало зусиль, терпіння та уважності,

оскільки вони розраховані на великий об'єм матеріалу документа, наприклад листування чи веб-сторінки. Пошук dtSearch дає змогу виявити синонімічний ряд, неточний пошук та фонетичний. Неточний пошук включає в себе елементи алгоритму відстані Левенштейна, знаходячи схожі слова, які починаються з однакової букви, тоді як фонетичний пошук більш схожий зі саундексом.

Всесвітня організація інтелектуальної власності пропонує свою систему пошуку ROMARIN на CD-дисках та онлайн, яка містить усі наявні міжнародні реєстрації торгових марок, а також посилання на dtSearch для перевірки даних за заявкою.

Висновки. Таким чином, для отримання найбільш точного та правдивого результату пошуку словесної схожості торгових марок зі ступенем змішання з іншими варто користуватися консолідованим алгоритмом пошуків, тобто поєднувати різні види пошуку та комбінувати варіанти. Жоден з алгоритмів пошуку чи пошукових систем не здатен виявити торгові позначення, у яких змінений буквенний ряд, наприклад, SIDALENE та SADINELE, а також у випадку, якщо відсутня частина слова (YOGODIP та YOCO).

Література

1. Про охорону прав на знаки для товарів і послуг : Закон України від 15 грудня 1993 р. № 3689-ХІІ // Відомості Верховної Ради України. – 1994. – № 7. – Ст. 36.
2. Державна служба інтелектуальної власності України [Електронний ресурс]. – Режим доступу : <http://sips.gov.ua/ua/systems.html>.
3. Кириченко І. Судова експертиза об'єктів права інтелектуальної власності : [навчальний посібник] / І. Кириченко. – К. : Ін-т інтел. власн. і права, 2005. – 172 с.
4. Всесвітня організація інтелектуальної власності [Електронний ресурс]. – Режим доступу : <http://www.wipo.int/classifications/vienna/en/>.
5. Brodda B. Corpus work with PC beta: a presentation / B. Brodda // Proceedings of the 13th conference on computational linguistics. – Morristown ; NJ : Association for Computational Linguistics, 1990. – Vol. 3. – 405 p.
6. Fall C. Searching trademark databases for verbal similarities / C. Fall // World Patent Information. – Switzerland : Published by Elsevier Ltd, 2005. – № 27. – P. 135–143.
7. Bugdahl V. Markenrecherchen – eine subjective Momentaufnahme / V. Bugdahl // MarkenR. – 2003. – № 5. – P. 169–180.
8. Соколов А. Методи нейромережевого розподіленого представлення та пошуку схожих символічних послідовностей в задачах класифікації на основі міркувань за прикладами : автореф. дис. ... канд. юрид. наук : спец. 05.13.23 «Системи та засоби штучного інтелекту» / А. Соколов ; Міжнародний науково-навчальний центр інформаційних технологій та систем. – К., 2008. – 21 с.
9. Бузикашвили Н. N-грамми в лингвистике / Н. Бузикашвили, Г. Крылова, Д. Самойлов // Методы и средства работы с документами. – М. : Диториал УРПС, 2000. – С. 91–130.
10. Zobel J. Finding approximate matches in large lexicons / J. Zobel, P. Dart // Software. Practice & Experience. – 1995. – № 25. – P. 331–345.
11. Zobel J. Phonetic String Matching: Lessons from Information Retrieval / J. Zobel, P. Dart // SIGIR : proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval. – New York, 1996. – P. 166–172.
12. Philips L. The double metaphone search algorithm / L. Philips // Users journal. – 2000. – Vol. 18(5). – P. 38–43.
13. Holmes D. Improving precision and recall for soundex retrieval / D. Holmes // Coding and Computing: proceedings of the International Conference on Information Technology. – Nevada : IEEE Computer Society, 2002. – P. 22–27.