

**УДК 004.414**

**А.В. Товпига, О.В. Маєвський**

Тернопільський національний технічний університет імені Івана Пулюя, Україна

## **ПОКАЗНИКИ ЕФЕКТИВНОСТІ ІНФОРМАЦІЙНОГО ПОШУКУ**

**A.V. Tovpyha, O.V. Majevskiy**

### **PERFORMANCE INDICATORS INFORMATION SEARCH**

Пошук інформації в Інтернет у міру збільшення об'єму розосередження її джерел стає все більш складним і трудомістким. При цьому критичним є добір інформації, яка відповідає запиту користувача.

Ступінь відповідності видачі пошукових результатів очікуванням користувача, який звернувся з пошуковим запитом, називається релевантністю. За ступенем релевантності видачі судять про ефективність роботи пошукової системи. Наскільки релевантна та чи інша інформація визначається пошуковим алгоритмом. У найпростішому наближенні релевантність сторінки визначається співвідношенням кількості вживання ключових слів, що входять в пошуковий запит, до загального обсягу ключових слів на сторінці [1]. Для кожної системи існує мінімальний відсоток входження пошукової фрази в текст, що говорить про достатню релевантність сторінки. В середньому це 5%. Менша кількість пошукових фраз може бути проігноровано пошукачем. Більша - може потрапити під обмеження спам-фільтра пошукової системи з усіма витікаючими наслідками [2].

Якщо під релевантністю розуміється формальна відповідність запиту інформації, що видається системою, то на практиці використовується інше, неформальне поняття – пертинентність. Для користувача пертинентність, співвідношення обсягу корисної для нього інформації до загального обсягу отриманої інформації, має вирішальне значення. Досягнення високої пертинентності – основне поле конкурентної боротьби сучасних пошукових систем. Саме для максимального задоволення інформаційних потреб користувачів мережеві інформаційно-пошукові системи сьогодні максимально інтелектуалізуються – отримали широке застосування технології і методи семантичних і нейронних мереж, Text Mining.

Існує багато способів оцінити наскільки добре документи, знайдені інформаційно-пошуковою системою, відповідають запиту. Щоб отримати числові значення пошуку використовується ряд показників: повнота, точність, втрата інформації, інформаційний шум, F-міра [3].

Для підрахунку коефіцієнтів повноти, точності та F-міри використовується формули (1), (2), (3) відповідно.

$$r = \frac{a}{a + c}, \quad (1)$$

де  $r$  – коефіцієнт повноти;

$a$  – релевантні документи, знайдені системою;

$c$  – релевантні документи, не знайдені системою.

$$p = \frac{a}{a + b}, \quad (2)$$

де  $p$  – коефіцієнт точності;

$b$  – не релевантні документи, знайдені системою.

$$F = \frac{2}{\frac{1}{p} + \frac{1}{r}}, \quad (3)$$

де  $F$  – коефіцієнт  $F$ -міри.

Втрата інформації розраховується як відношення числа не знайдених релевантних документів до загального числа релевантних документів в базі (4) [4].

$$BI = \frac{C}{A+C} \quad (4)$$

Інформаційний шум – відношення знайдених не релевантних документів до загального числа знайдених документів системою (4).

$$IШ = \frac{B}{A+B} \quad (5)$$

Отже для оцінки ефективності інформаційного пошуку в мережі Інтернет і семантичній «павутині» варто використовувати показники ефективності – точність, повноту та  $F$ -міру. В основу оцінки релевантності знайдених документів покладено твердження, що весь пошуковий простір відомий і тому перетин знайдених ресурсів і створеного пошукового простору дасть обсяг релевантних документів (див. рисунок 1).

Знайдені ресурси

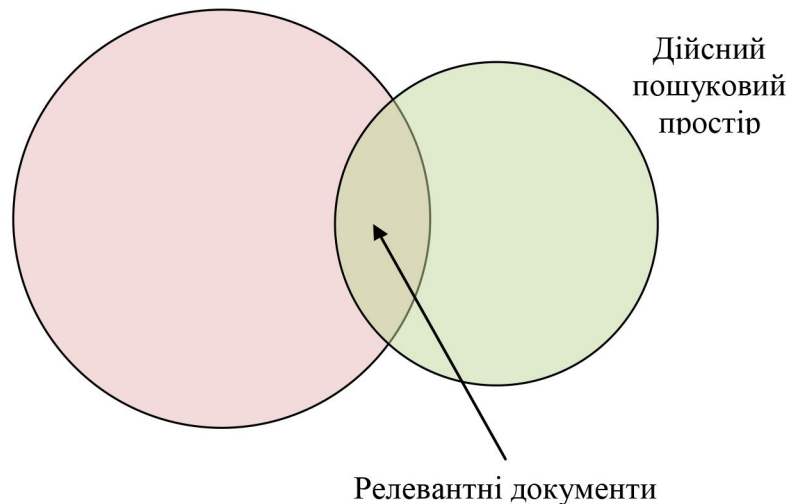


Рис. 1. Графічне зображення вибору релевантних документів

### Література

1. Semantic Web Challenge [Електронний ресурс] - Режим доступу до ресурсу: <http://challenge.semanticweb.org/>.
2. Д.А.Лапко. Аналіз ефективності інформаційного пошуку в мережі інтернет з використанням метаданих - страница 2 / Д.А. Лапко // Моя бібліотека. – Режим доступу: [http://www.vuzlib.com.ua/articles/book/20299-Anal%D1%96z\\_efektivnost%D1%96\\_%D1%96nforma/2.html](http://www.vuzlib.com.ua/articles/book/20299-Anal%D1%96z_efektivnost%D1%96_%D1%96nforma/2.html). – Дата доступу: жовтень 2015 року. – Заголовок з екрану.
3. Рогушина Ю.В. Средства интеллектуализации поиска информационных ресурсов в сети Интернет / Ю.В. Рогушина // Материалы науч.-практ. конф «Наука і освіта 2003». – Дніпропетровськ: 2003 Том 28. - С. 19 - 20.
4. Інформаційний пошук // Знаймо. – Режим доступу: [http://znaimo.com.ua/Інформаційний\\_пошук](http://znaimo.com.ua/Інформаційний_пошук). – Дата доступу: жовтень 2015 року. – Заголовок з екрану.