

УДК 004.01(082)

Воротін О.С., Дехтяренко Д.Т.

Національний авіаційний університет

Фаховий коледж інформаційних технологій та землевпорядкування

Національного авіаційного університету

МОДЕЛЬ ІНФОРМАЦІЙНОГО ПОШУКУ

В даний час інформаційні ресурси тільки веб-простору складають понад 20 мільярдів документів, до яких можливий вільний доступ будь-якого користувача. Природно, для того, щоб знайти необхідну інформацію і цієї найбільшої розподіленої повнотекстової бази даних необхідно використовувати найпотужніші ПС [1]. В основу роботи всіх ПС покладено спеціальні алгоритми, які є модифікаціями основних підходів – моделей пошуку.

Модель інформаційного пошуку має три ключових аспекти.

1) Формат представлення документа. Документи можуть містити тексти на природній або формалізованій мові, зображення, звукову інформацію тощо. 2) Формат представлення запиту. Використовується мова пошукових запитів, синтаксис якої змінюється в рамках різних систем. 3) Функція відповідності документа запиту. Ступінь відповідності запита і знайденого документа (релевантність) – суб'єктивне поняття, оскільки результати пошуку, що задовольняють одного користувача, можуть не задовольняти іншого.

В основу традиційних методів покладено три головні підходи, перший з яких базується на теорії множин (булева модель), другий – на векторній алгебрі (векторно-просторова модель), а третій – на теорії ймовірностей (імовірнісна модель). Ці підходи можуть застосовуватися на практиці і в канонічному вигляді, проте у них є спільний недолік, обумовлений припущенням, що зміст документа визначається безліччю слів і стійких словосполучень – термів (англ.– Terms), які входять в нього без урахування взаємозв'язків, і, більше того, вважаються незалежними. Таке припущення веде до втрати змістовних відтінків, проте воно дозволяє реалізувати пошук і групування документів за формальними ознаками.

Використані джерела

1. *Димлюк О.М.* Пошукові системи мережі Internet: технології та засоби. Міжнародна науково-технічна конференція «Інтелектуальні технології лінгвістичного аналізу». Тези доповідей. – К.: НАУ, 2008. – С.25.