УДК 004.01(082)

А.С. Шевченко

*Національний авіаційний університет*

СТРАТЕГІЇ ІНФОРМАЦІЙНОГО ПОШУКУ

Стратегії інформаційного пошуку визначають ступінь подібності документів, що розглядаються, до пошукового запиту. Ступінь подібності визначається згідно робочої гіпотези: чим частіше пошуковий термін зустрічається в документі, тим більше «відповідним» є цей документ до пошукового запиту.

Стратегії інформаційного пошуку розробляються не тільки для визначення відповідності, але і для вирішення проблем, які пов’язані з неоднозначністю мови – один і той самий термін може позначати різні концепти (ключ в механіці означає зовсім не те, що в шифруванні), один і той же концепт може позначатись різними термінами. Стратегія інформаційного пошуку це алгоритм, який, переглядаючи набір документів (Д1, ..., Дn), встановлює їх відповідність до пошукового запиту (ПЗ). Оскільки пошуковий термін зустрічається в документах різну кількість раз, можна говорити про різну ступінь відповідності до пошукового запиту. Алгоритм обчислює коефіцієнт відповідності) (КВ) для кожного документу КВ(ПЗ, Дi), де 1 ≤ i ≤ n.

Існують такі стратегії інформаційного пошуку:

* з використанням векторно-просторового представлення;
* пошук імовірності появи пошукового терміну в документі;
* з побудовою мовної моделі для кожного документу;
* з побудовою мережі припущень, яка використовується для встановлення відповідності документу до пошукового запиту;
* з булевим індексуванням, коли кожному пошуковому терміну присвоюється своя «вага», що потім враховується при побудові впорядкованих списків документів;
* з використанням не проявленого семантичного індексування;
* з побудовою нейромереж;
* з використанням продуктивних алгоритмів, коли початковий пошуковий запит «еволюційно» видозмінюється;

з використанням нечітких множин, коли документу ставиться у відповідність нечітка множина.