

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНТЕЛЛЕКТУАЛЬНОГО ТЕМАТИЧЕСКОГО ПОИСКА

Чередниченко О.Ю., Норбутаев А.А.
*Национальный Технический Университет
«Харьковский политехнический институт»,
г. Харьков*

Современные условия века информационных технологий диктуют необходимость исследований в области применения информации глобальных сетей для решения различных экономических, маркетинговых, образовательных и т.п. задач. Веб-пространство накопило огромные объемы информации, которые постоянно увеличиваются. Современные пользователи компьютерных технологий все чаще обращаются к глобальной сети, решая как бытовые, так и коммерческие задачи. Большинство этих задач связаны с необходимостью поиска некоторой информации на заданную тему. Проблема заключается в том, что, начиная тематический поиск, пользователь не может четко сформулировать информационный запрос, не знает как определить пространство поиска и, часто, не может сформулировать цели поиска. В процессе информационного поиска поисковый запрос постоянно уточняется. Однако, чтобы получить результат пользователю приходится просматривать сотни веб-страниц, которые оказываются нерелевантными теме поиска. Все это обуславливает необходимость исследований для теоретического обоснования и практической реализации программного обеспечения информационных систем тематического поиска.

Проведенный анализ показал, что проблемами тематического поиска занимаются многие исследователи. В данной работе предлагается рассматривать процесс тематического поиска информации в сети интернет как интеллектуальный процесс. Для моделирования процессов интеллектуальной деятельности человека предлагается применить подход на основе метода компараторной идентификации, который хорошо зарекомендовал себя в теории интеллекта.

Суть метода заключается в том, что сложный интеллектуальный процесс представляется как суперпозиция конечных детерминированных предикатных функций, которые реализуют функции эквивалентности и предпочтения. Просматривая страницу, пользователь всегда однозначно может сказать соответствует она или нет теме поиска, а также может сравнить две страницы в смысле близости теме. Это дает основание для разработки оценочной функции. Предлагается рассматривать следующие типы функций: оценка перспективности веб-страницы, оценка релевантности теме и оценка семантической близости. Реализации данных функций посвящено дальнейшее исследование.