

## ИССЛЕДОВАНИЕ МЕТОДОВ ПОВЫШЕНИЯ РЕЛЕВАНТНОСТИ РЕЗУЛЬТАТОВ ПОИСКОВЫХ ЗАПРОСОВ

Кольченко М.В.

*Харьковский национальный университет радиоэлектроники, г. Харьков*

Интернет является ценным источником информации с объемами, насчитывающими сотни миллионов сайтов и триллионы страниц. Однако информация в Интернете неупорядочена и неоднородна, и связующим звеном между пользователем, который хочет найти нужную информацию и владельцем сайта, желающего получить заинтересованного пользователя, являются поисковые системы.

Несмотря на достоинства современных поисковых систем, пользователи по разным причинам не ощущают удовлетворения от результатов поиска в сравнительно новой информационной среде, которой является Интернет.

Разработчики постоянно совершенствуют поисковые системы, стремясь сделать их алгоритмы максимально приближенными и похожими на то, как мыслят и действуют живые пользователи. Несмотря на это, понятие релевантного результата для компьютера и для пользователя отличается.

Предлагается использовать универсальный и не зависящий от конкретной поисковой машины метод определения релевантности ресурса запросу пользователя – отслеживание поведения пользователя в процессе его взаимодействия с веб-ресурсом.

Можно выделить ряд типичных действий (поведенческих признаков) пользователей, основываясь на которых, можно сделать вывод об интересе пользователя к конкретному веб-ресурсу. Эти признаки можно разделить на две группы: явные и неявные. Явные поведенческие признаки это: сохранение веб-ресурса в закладки или избранное, голосование, написание комментариев, поиск по сайту, подписка на рассылку или RSS. К неявным поведенческим признакам можно отнести: количество и время просмотра страниц на сайте, количество возвратов, переходы со страницы по внешним или внутренним ссылкам, активность на сайте и другие.

Дальнейшей задачей является разработка методики расчета оценки релевантности Интернет-ресурсов (определение набора параметров, отражающих интерес пользователя к веб-страницам) и разработка алгоритма оценки релевантности ресурсов, основанного на интересах пользователя.