

Санкт-Петербургский Государственный Политехнический Университет
Институт прикладной математики и механики
Кафедра прикладной математики

Диссертация допущена к защите
Зав. кафедрой

_____ В.Е.Клавдиев
" _ " _____

ДИССЕРТАЦИЯ
на соискание степени МАГИСТРА

Тема: *метод ранжирования разнородных результатов поиска*

Направление: 010400 - Прикладная математика и информатика
Магистерская программа: системное программирование

Выполнил студент гр. 63601/2 _____ Толмачев А.С.

Руководитель _____ к.ф.-м.н., доцент Иванков А.А.

Консультанты:

по вопросам информационного поиска _____ к.ф.-м.н. Кураленок И.Е.

по вопросам охраны труда _____ к.т.н., доцент Монашков В.В.

Санкт-Петербург
2015

Содержание

Введение	2
1 Обзор литературы	3
2 Постановка задачи	4
Заключение	5

Введение

Глава 1

Обзор литературы

Глава 2

Постановка задачі

Заключение

В данной работе предложен новый метод ранжирования разнородных результатов поиска. Главная отличительная особенность метода состоит в том, что результаты поиска располагаются исходя из соображений максимизации релевантности всей поисковой выдачи в целом, а не в соответствии с релевантностями отдельных результатов. Благодаря этому метод является универсальным – он может быть применен к различным видам поисковых результатов и для разных моделей поисковой выдачи. Также переход от рассмотрения поисковых результатов по отдельности к рассмотрению выдачи в целом позволяет естественным образом учитывать зависимости и отношения между разными типами результатов. (Еще преимущества?)

Предложенный метод был реализован и применен для встраивания 31 типа специализированных результатов вертикальных поисковых источников в мобильную поисковую выдачу системы Яндекс. Использовались специализированные результаты поиска по картинкам, видео, мобильным приложениям, поиска товаров, новостей, погоды, результаты гео-поиска и других сервисов компании Яндекс. Была проведена оценка качества работы метода [ref] и сравнение с текущим используемым методом встраивания специализированных результатов [ref] по метрикам, основанным на ассессорских оценках: по точности и полноте показа специализированных результатов и метрике *pfound* [ref] (+ **online-метрики?**). Сравнение показало улучшение точности показа специализированных результатов на 21.22% при снижении полноты на 29.21% и прирост качества по метрике *pfound* на 0.27%. (TODO: уточнить результаты)

В ходе реализации метода и встраивания его в поисковую систему также была решена задача эффективного нахождения аргумента максимизации функции, представляющей собой ансамбль решающих деревьев специального вида (oblivious decision

trees), и нахождения заданного числа кандидатов в аргументы максимизации при наличии частично вычисленного вектора признаков [ref]. Решение этой задачи позволяет избежать задания запросов к тем поисковым источникам, результаты которых будут заведомо нерелевантны заданному поисковому запросу. Также следует отметить, что решение данной задачи имеет самостоятельную ценность, и может быть применено не только для реализации предложенного метода ранжирования разнородных результатов поиска, но и в других задачах.