

## **Алгоритм работы поискового модуля в системах поддержки принятия решений**

Зеленков П.В., Ковалев И.В., Карасева М.В., Рогов С.С.

Сибирский Государственный Аэрокосмический Университет

В настоящее время активно развивается направление разработки и внедрения корпоративных информационно-управляющих систем. В рамках данных систем наиболее интересными технологиями являются технологии поддержки принятия решений, однако в них недостаточно уделено внимание проблеме поиска необходимой информации для лиц принимающих решения. Авторы предлагают использовать следующий алгоритм, состоящий из ряда процессов.

Первый процесс при выполнении поисковой процедуре – это процесс первичного анализа проблемы. На данном этапе происходит задание начальных параметров поиска и инициализация поисковой процедуры. Также производится выбор одного из четырех возможных путей выполнения процедуры поиска (поиск без использования тезаурусов, поиск с использованием тезауруса, поиск с использованием частотного тезауруса и поиск по уже анализируемому ранее запросу).

После того как будет произведена поисковая процедура по одному из возможных вариантов, необходимо обработать полученные данные для определения релевантности и ранжирования информации. Следующим шагом предлагаемого алгоритма является просмотр полученных и обработанных данных лицом принимающим решение.

Теперь рассмотрим каждый из процессов более подробно.

Первый процесс – это процесс первичного анализа проблемы. В начале, необходимо указать тип проведения поисковой процедуры. Здесь можно выделить следующие возможные типы проведения поиска: по заданному информационному ресурсу корпоративной сети, по информационному

ресурсу в Internet, метапоиск по заданным информационным ресурсам корпоративной сети, либо по всем ресурсам корпоративной сети, метапоиск по заданным информационным ресурсам Internet, метапоиск по заданным поисковым ресурсам Internet, смешанный метапоисковый алгоритм, как по ресурсам корпоративной сети, так и в Internet.

Далее необходимо определить предметную область поиска. На данном этапе необходимо указать тезаурус или ключевые слова предметной области (необязательно имеющие отношение к конкретной проблеме из предметной области) в рамках которой необходимо будет проводить поисковую процедуру.

Далее происходит указание обрабатываемых языков. Это очень важная характеристика, так как она связана с мультилингвистичностью, как поисковой строки, так и ответа системы.

Далее необходимо указать информационные ресурсы в рамках которых будет проведена поисковая процедура. Как говорилось выше, данная функция генерирует список ресурсов динамически в зависимости от выбора типа поиска.

Следующий шаг – это формирование поисковой строки. Данный шаг может быть выполнен с помощью тезаурусов или вручную.

И наконец, после того как пользователь определился с проблемой и задал все начальные характеристики, необходимо выполнять поисковую процедуру по четырем алгоритмам.

Первый алгоритм поисковой процедуры возникает в случае возникновения новой ситуации, по которой ранее поиск не проводился.

Данный режим работы поисковой системы наиболее прост в исполнении, но дает наихудшие результаты. Здесь производится поиск по сформированной вручную, поисковой строке.

Второй и третий возможные алгоритмы работы поисковой процедуры относительно похожи (поиск с использованием обычного тезауруса и частотного тезауруса). Важной частью работы алгоритма является

использование мультилингвистичных тезаурусов. Необходимо отметить, что при одноязычной поисковой строке выбран мультилингвистический поиск, система автоматически переведет одноязычную строку поиска в мультилингвистическую, используя тезаурус. Так же система проверит наличие мультилингвистичности термов заданной поисковой строки.

Последний режим работы поисковой процедуры - проведение поисковой процедуры уже решаемой ранее проблемы. Здесь сыграет существенную роль уже просмотренные документы ранее. То есть если лицо принимающее решение ранее просматривал документ представленный системой, то его релевантность будет выше и при ранжировании данные документы будут на первом месте.

После проведения поисковой процедуры необходимо провести обработку полученного множества данных. Здесь можно провести дополнительную проверку уровня релевантности каждого документа. Исходя из показателя уровня релевантности, происходит ранжирование документа.

Таким образом, показаны основные этапы проведения поиска информации в корпоративных системах.

---

Работа представлена на заочную электронную конференцию «Новые информационные технологии и системы, 15-20 декабря 2007г.». Поступила в редакцию 05.02.2008г.