

PEDAGOGIK MAHORAT

MS
2022



MUNDARIJA

№	Familiya I.Sh.	Mavzu	Bet
1.	БАКАЕВ Илхом Иззатович, ЭШАНКУЛОВ Хамза Илхомович	Формирование механизма поиска с применением алгоритмов полнотекстового поиска	7
2.	ЖАЛОЛОВ Озоджон Исомидинович, БАРНОЕВА Зубайда Эркин кизи, ИСОМИДДИНОВ Бекзоджон Озоджон угли	Методы построения оптимальной весовой квадратурной формулы типа эрмита в пространстве периодических функций Соболева $\tilde{W}_2^{(m)}(T_1)$	14
3.	ШАФИЕВ Турсун Рустамович, САЛИМОВ Рузбек Насим угли	Алгоритм сопоставления отпечатков пальцев	20
4.	JUMAYEV Jo'ra, ISMATOVA Kamola Otabek qizi	Transport masalasini kompyuterli modellash	27
5.	RUSTAMOV Hakim Sharipovich, QURBONOV Suhrob Bekro'latovich	Zamonaviy axborot-kommunikatsiya texnologiyalaridan foydalanish ta'lim samaradorligining asosiy omili	32
6.	ZARIPOVA Gulbahor Kamilovna, HAZRATOVA Roila Zainiddinovna	Development of professional competence of specialists in the training of teachers in digital and information technologies in our society	36
7.	XAZRATOV Fazliddin Xikmatovich, RUFATOV Jo'rabek Zafar o'g'li	Data mining qo'llash sohasi. Prognozlash va vizualizatsiya masalalarini hal etish	43
8.	ЖАЛОЛОВ Озоджон Исомидинович, НАСРИДДИНОВА Халима Фарход кизи, РАСУЛОВА Камола Хаким кизи	Методы построения оптимальных по порядку сходимости кубатурных формул типа эрмита в пространстве соболева	50
9.	АТАЕВА Гулсина Исроиловна, МАХМАДИЕВ Хасан	Роль искусственного интеллекта в образовании	57
10.	TURDIEVA Gavhar Saidovna	Kredit modul tizimida talabalarning ilmiy-tadqiqot ishlari - mustaqil faoliyatning eng yuqori shakli sifatida	62
11.	TURDIEVA Gavhar Saidovna, DJURAYEVA Salomat Nabiyevna	Ta'lim jarayonida stem-texnologiya-talabalarning loyihalash faoliyatini rivojlanish vositasi sifatida	68
12.	ШАФИЕВ Турсун Рустамович, ЭШОНКУЛОВ Шахзод Равшанович	Аутентификация личности на мобильных устройствах с использованием проверки	73
13.	IMOMOVA Shafolat Mahmudovna	Matematikani o'qitishda matematik tizimlardan foydalanish	77
14.	IMOMOVA Shafolat Mahmudovna, BOTIROVA Nigora Qoyirovna	Google classroom - "virtual sinf" texnologiyasi	81
15.	JUMAYEV Jo'ra, SHAMSIYEVA Nigora Rafiq Qizi	Chiziqli dasturlash masalasini simpleks usulda yechishning kompyuterli modeli	86
16.	ИСМОЙЛОВА Махсума Нарзикуловна, НАМОЗОВА Нигина Шермат кизи	Методы и дидактические задачи на основе мобильных технологий обучения	91
17.	YADGAROVA Lola Djalolovna, ERGASHEVA Sarvinoz Bahodurovna	Innovative approach: project-based learning the organization of the educational process in higher educational institutions	96

БАКАЕВ Илхом
Иззатович

ЭШАНКУЛОВ Хамза
Илхомович

Доцент кафедры информационных систем
и цифровых технологий Бухарского
государственного университета

Доцент кафедры «Математика и технология
программирования» Бухарского
государственного университета

ФОРМИРОВАНИЕ МЕХАНИЗМА ПОИСКА С ПРИМЕНЕНИЕМ АЛГОРИТМОВ ПОЛНОТЕКСТОВОГО ПОИСКА

В статье разработана методология организации поиска с применением алгоритмов полнотекстового поиска. Построено представление о принципе работы создаваемой поисковой машины. Определены основные этапы организации внутреннего поиска, а также сделан обзор для принятия решения о применении специальных разработанных проектов, с помощью этих поисковых машин можно реализовать поиск на сайте. Приведены примеры при организации поиска с учётом языковой морфологии и использования двух специально созданных морфологических модулей – стеммера и лемматизатора.

Ключевые слова: стеммер, лемматизатор, полнотекстовый поиск, алгоритм индексации.

TO'LIQ MATNLI QIDIRISH ALGORITMLARIDAN FOYDALANIB QIDIRUV MEKANIZMINI SHAKLLANTIRISH

Maqolada to'liq matnli qidiruv algoritmlari yordamida qidiruvni tashkil qilish metodologiyasi ishlab chiqilgan. Yaratilgan qidiruv tizimining ishlash prinsipi haqida g'oyalar qurilgan.

Ichki qidiruvni tashkil etishning asosiy bosqichlari aniqlanadi va ushbu qidiruv tizimlari yordamida saytda qidiruvni amalga oshirish uchun ishlatilishi mumkin bo'lgan maxsus ishlab chiqilgan loyihalardan foydalanish to'g'risida qaror qabul qilish uchun ko'rib chiqiladi. Til morfologiyasini hisobga olgan holda qidiruvni tashkil qilishda va ikkita maxsus tayyorlangan morfologik modullardan - stemmer va lemmatizatoridan foydalanishga misollar keltirilgan.

Kalit so'zlar: stemmer, lemmatizator, to'liq matnli qidiruv, indekslash algoritmi.

FORMING A SEARCH MECHANISM USING FULL-TEXT SEARCH ALGORITHMS

The article develops a methodology for organizing search using full-text search algorithms. Ideas about the principle of operation of the created search engine are built.

The main stages of organizing an internal search are determined, and a review is made to decide on the use of special developed projects that, with the help of these search engines, can be used to implement a search on the site. Examples are given when organizing a search taking into account language morphology and using two specially made morphological modules - a stemmer and a lemmatizer.

Keywords: stemmer, lemmatizer, full-text search, indexing algorithm.

Введение. С увеличением спроса пользователей на потребность информации, день за днём растёт количество сайтов, которые специализируется на каком-либо устном направлении. При создании сайта все владельцы в первую очередь задумывались о вопросе организации поиска на сайте, потому что для сайтов с большим контентом и многотысячной аудиторией поиск является основной неотъемлемой частью в процессе обмена информации.

Для поиска данных из сайта имеется множество алгоритмов. До создания поисковых машин, работающих на основе машинного обучения, поиск в основном был организован методом прямого поиска искомого запроса из базы данных сайта. Это было очень трудоёмкой задачей, и в большинстве случаев поисковая выдача возвращала пустые значения. Тем не менее специалисты разрабатывали разные алгоритмы и методы для формирования поиска данных из сайта.

В решение проблемы задачи организация поиска на сайте с применением алгоритмов полнотекстового поиска и усовершенствования существующих поисковых алгоритмов большой вклад внесли такие учёные, как Zhitao Guan, Hong Zhong, Alexis Kaporis, Tessa L. Crume, Кизянов А. Ф, Соченков И.В., Суворов Р.Е и другие.

В частности, в статье авторов [1] рассмотрена задача эффективного поиска данных в Web приложении, и проанализирован механизм полнотекстового поиска, позволяющий реализовывать удобные средства поиска интересующей информации по содержимому документов. Оценена эффективность выполнения поисковых запросов в СУБД MySQL, PostgreSQL и Oracle.

В статье [2] обсуждён метод поиска по ключам, основанный на известных механизмах образного мышления. Обработка запроса к базе данных начиналась с построения компьютерного образа искомой записи в виде совокупности её ключевых характеристик. Для этого образ искомой записи был организован как форма с представлением результатом исследования, хорошо зарекомендовавшая себя среди экспериментов на основе гистограмм распределения измеренных данных. Столбцы гистограммы с максимальными значениями высоты рассмотрены как наиболее вероятные номера искомых записей.

В работе [3] разработана специальная поисковая машина, работающая на основе алгоритма Sphinx. Анализируя особенности текстового поиска, автор предлагает учитывать морфологические основы языка. Также в работе представлены установки и конфигурации поисковой машины, работающей на основе алгоритма Sphinx. Рассматриваются «базовые» концепции и методы индексирования, которые необходимо знать и использовать ежедневно при работе со Sphinx.

В статье [4] рассматриваются алгоритмы последовательного и логарифмического поиска (наихудший и средний случай), даётся их анализ, оцениваются эффективность и время выполнения. Приведены образцы выполнения заданий, демонстрационные примеры.

Диссертация автора [5] сфокусирована на проблеме получения текста, допускающей ошибки, также называемой «приблизительным» сопоставлением строк. Изучена проблема поиска шаблона в тексте, где шаблон и текст могут иметь «ошибки». Решение этих проблем привлекало много внимания в последние годы из-за его применения во многих областях, таких как информационный поиск, вычислительная биология и обработка сигналов. Основной целью данной работы является разработка и анализ новых алгоритмов для решения проблемы поиска текста из источника в различных условиях, а также лучшее понимание самой проблемы и её статистического поведения. В конечном итоге автором разработан алгоритм поиска с поддержкой исправления “ошибочного” запроса.

В работе [6] сделан обзор основных подходов и методов, применяемых для решения ряда задач информационного поиска по текстовым данным. Данный обзор не призван охватить и рассмотреть абсолютное большинство задач информационного поиска, однако, как надеется автор, он будет полезен в качестве общего освещения методов поиска, применяемых в современных информационно-поисковых системах.

В статье [7] рассматривается задача полнотекстового поиска в корпусе текстов. Автор для решения предлагает использовать алгоритм с обучением, сочетающий поиск по словосочетаниям с анализом данных о поведении пользователей. В конце приводятся результаты эксперимента, подтверждающие эффективность данного алгоритма.

В работе автора [8] рассматриваются задачи поиска фраз и наборов слов в большом объёме текстов. Основным результатом поиска является получение списка документов, содержащих заданные слова, при этом документы, где слова располагаются ближе друг к другу, считаются более релевантными. Поскольку эта задача требует сохранения в индексе информации о каждом вхождении каждого слова в текстах, запросы, включающие часто встречающиеся слова, требуют для своего выполнения длительного времени. В некоторых поисковых системах предлагается ввести список стоп слов, которые не учитываются при поиске, но этот подход снижает качество поиска. В данной работе при поиске обрабатываются все слова, и применяются дополнительные индексы. С помощью дополнительных индексов время выполнения поискового запроса, включающего часто встречающиеся слова, может быть снижено в десятки раз. Разработан новый вид индекса с трёхкомпонентными ключами. Приведены алгоритмы поиска и результаты экспериментов поиска в сравнении с обычными индексами. Эксперименты, проведённые автором, показывают, что при применении разработанных индексов для определённого класса запросов, состоящих из самых часто встречающихся слов, скорость поиска возрастает более, чем в 90 раз.

В статье [9] исследование посвящено разработке структур, данных и алгоритмов информационного поиска. Сделан анализ описание разработанных структур данных для

представления поисковых индексов в информационно-аналитической системе. Предлагаемые структуры данных ориентированы на представление лингвистической информации текстов, включая синтаксическую и семантическую. Рассмотрены алгоритмы построения поисковых индексов и проведены практические примеры по индексации страниц.

В статье авторов [10] приводится анализ алгоритмов, которые позволяют пользователям осуществлять поисковые операции в интеллектуальных системах, прогнозировать их мысли и находить ключевые слова, которые они ищут. Также рассказывается о том, как эти алгоритмы можно использовать в интеллектуальных библиотечных системах, что позволяет экономить время пользователей, затрачиваемое на поиск.

В статье авторов [11] предлагается механизм извлечения документов с использованием оптимизированных алгоритмов основанной на оптимизации Monarch Butterfly FireFly (MB-FF) с интеграцией оптимизации Monarch Butterfly (MBO) и алгоритма Firefly (FA). Ключевые слова из документов идентифицируются из предварительно обработанного документа, который предварительно обрабатывается с использованием основ и удаления стоп-слов. Термин «частотно-обратная частота документов» (TF-IDF) используется при извлечении ключевых слов, а понятие голоэнтропии - при выборе значимых ключевых слов. Выбранные ключевые слова обеспечивают поиск соответствующих документов, которые первоначально обрабатываются с помощью кластерной индексации с использованием светлячка на основе оптимизации Monarch Butterfly (MB-FF), за которым следует двухуровневое соответствие мод-Бхаттачарья. Производительность алгоритма MB-FF в механизме поиска документов оценивается с использованием точности, и F-измерения.

В работе [12] представлен алгоритм поиска документов по заданной теме на основе выбранной справочной коллекции документов с созданием контекстно-семантического графа для визуализации тем в результатах поиска. Алгоритм основан на интеграции набора вероятностных, энтропийных и семантических маркеров для извлечения взвешенных ключевых слов и комбинаций слов, которые описывают данную тему. Результаты тестов демонстрируют среднюю точность 99% и отзыв 84% при экспертном отборе документов. Также авторами разработан специальный подход к построению графа на основе алгоритмов, извлекающих ключевые фразы с весами. Это даёт возможность продемонстрировать структуру под тем в больших коллекциях документов в виде компактных графов.

В работе [13] авторы представили нечёткий алгоритм для улучшения поиска документов, оптимизированный для конкретных сценариев, т.е. в поиск документов по не точно совпадающим словосочетаниям, если во множественном числе есть такие слова, или синонимам. Разработанный алгоритм применён в реальном времени, и результаты сравнены с результатами справочного менеджера Менделеев.

В работе [14] автор предлагает новый расширенный метод поиска совместного цитирования, который представляет собой поиск документов на основе графа в сети совместного цитирования, содержащей информацию о контексте цитирования. Предлагаемый алгоритм поиска расширяет сферу охвата целевых документов за счёт повторного распространения взаимосвязи. В частности, этот метод поиска представляет собой комбинацию алгоритмов: алгоритма на основе графов для вычисления показателя сходства в сложной сети и включения контекстов совместного цитирования в процесс вычисления показателей сходства для уменьшения негативных последствий растущего количества неактуальных документов. Чтобы оценить эффективность предложенного алгоритма поиска, предложены 10 методов (пять репрезентативных алгоритмов на основе графов, применённых к сетям совместного цитирования), которые сравниваются с двумя типами базовых показателей (традиционный поиск с совместным цитированием / без контекстов). Результаты эксперимента показали, что оценки нормализованного дисконтированного совокупного усиления предлагаемых методов с использованием контекстов совместного цитирования имели наилучшие результаты.

В статье [15] рассматривается возможность автоматизации процесса индексирования и поиска текста на естественном языке с применением самообучающегося поискового индекса. Автором предложена структура электронной библиотеки, удобная для организации поиска. Описан автоматический метод полнотекстового индексирования текста на естественном языке. Показано, что, при наличии определённых свойств языка, предложенный метод индексирования может быть

эффективен. Для предложенного метода индексирования описан алгоритм сопоставления проиндексированного текста запросу.

Подробный анализ научных работ, связанных с проблемой организации поиска с применением алгоритмов полнотекстового поиска показало, что организации поиска с применением выше указанных алгоритмов ещё не применено для узбекского языка. Но, тем не менее, на сегодняшний день у всех крупных поисковиков есть API, в котором есть уже готовые алгоритмы для организации полнотекстового поиска. В качестве примера можно взять поисковики Google или Yandex, и с помощью их API можно получить многие преимущества, такие как мощный морфологический анализ запроса, распознавание неверного раскладки клавиатуры, исправление опечаток и ошибок запроса. Конечно, эти преимущества имеют очень хорошие качества, но при применении этих API имеются некоторые нюансы и недостатки:

- Такой поиск не интегрируется в структуре сайта, так как он внешний, владелец сайта не может указать поисковику, какие данные очень важные, а какие можно пропустить.

- Проблема с индексированием (индексация сайта – это чтение роботами поисковых систем сайта, и внесение его в базу поисковика). Содержимое сайтов индексируется поисковыми роботами только в определённых интервалах, которые зависят от выбранного поисковика. Если на сайте будет осуществлено обновление, тогда владелец сайта должен дожидаться момента повторной индексации и только после этого его обновления станут доступными в поисковике.

1. Постановка задачи.

Исходя из вышеуказанного, возникает необходимость создания поискового механизма для внутреннего поиска поддерживающих агглютинативных языков типа узбекских языка. Для этого надо определить основные этапы организации внутреннего поиска, а также, нужно принять решения о применении специальных разработанных проектов, с помощью этих поисковых машин можно реализовать поиск на сайте. В целом, для организации внутреннего поиска определим следующие этапы:

2. Организация поиска с учётом языковой морфологии.

Данный способ позволяет находить все вхождения ключевого слова в строку текста, правильно обнаруживая любые грамматические формы искомого слова. В качестве примера можно взять слова «книг», «книге», «книги» - это одно формы одного того же слова «книга», которые что нужно учитывать в поисковом процессе. Одним из преимуществ данной цели является приведение каждого слова из поискового запроса в слова содержимого сайта к базовой форме.

Для эффективного определения ключевого слова заданного запроса и слов в обрабатываемом тексте надо использовать два специально сделанных морфологических модуля, называемых стеммером и лемматизатором.

3. Возможность указать контекст поиска, то есть возможность самостоятельно выбрать содержимое (контента) сайта, в пределах которого будет работать алгоритм, а также определить значимость контекста. Осуществлять поиск, опираясь на контекст, — значит, придерживаться установившегося в поиске уровня абстракции и использовать понятия заданного в нём семантического поля. Потерять контекст в поиске означает перестать понимать то, на что опирается пользователь, или интерпретировать его мысль в ином смысле, исходя из заданного в поиске семантического поля понятий. Например, рассмотрим поиск фрагмента текста из теологических данных. Предполагается, что поисковой запрос чаще всего будет содержать название искомой книги, поэтому, поиск по названию книг будет иметь наивысший приоритет. В качестве других свойств можно выделить поиск по свойствам книги, затем поиск по описанию.

4. Индексирование содержимого сайта. Для разового поиска данных этот пункт является необязательным. Но при формировании большого числа запросов, сервер принимая каждого соединения, инициализирует поисковой движок. Поэтому будет сложно прогнозировать о потребностях требуемых ресурсов и времени. Для этого, есть технология индексации. Индексирования сайта выполняется только при обновлении или добавления контента, а поиск выполняется только по индексу, а содержимое здесь не имеет значения.

5. Ранжирование. Механизм сортировки результатов поиска выполняется на основе оценки значимости найденных данных, а сама последовательность называется *ранжированием*. В данном алгоритме считается количество слов в документах (страницах) и потом, определяя их значения,

записывается в специальную базу. Также для каждой разновидности контента сайта при индексировании задаётся специально определённый коэффициент, который будет иметь значительную роль при формировании позиции в поисковой выдаче [10].

6. Метод решения

Для реализации вышеуказанных определений нам надо сформировать принцип работы создаваемой поисковой машины (рис 1.).

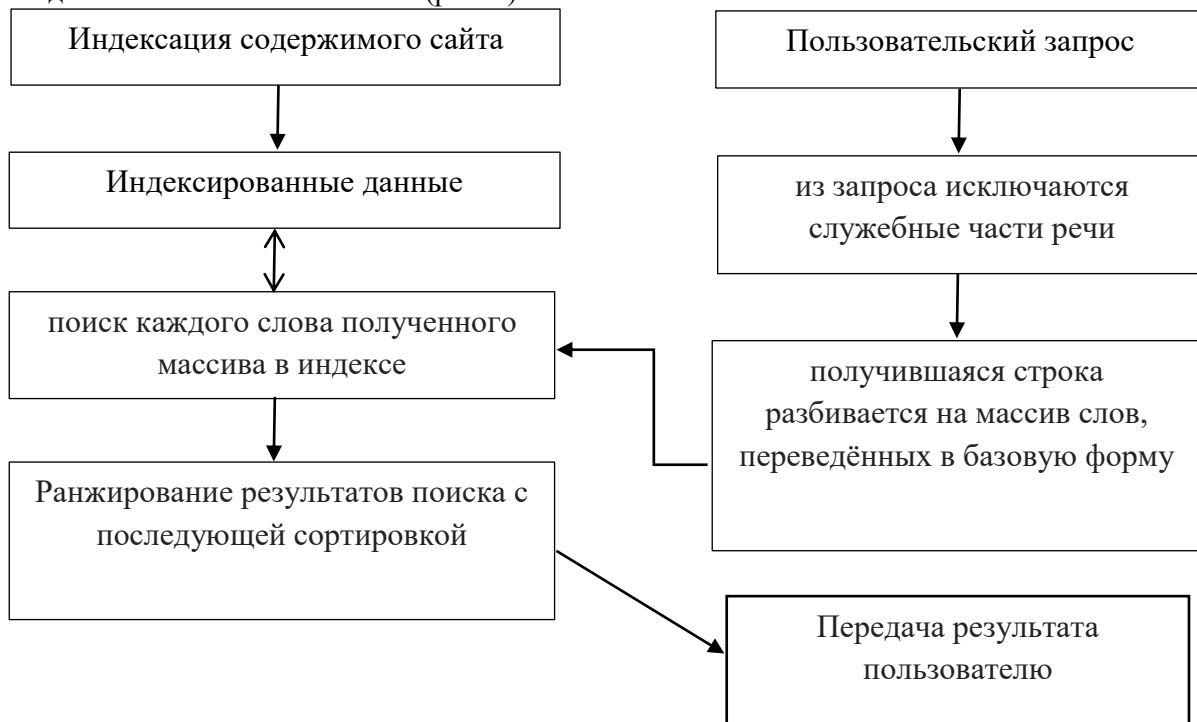


Рисунок 1. Принцип работы создаваемой поисковой машины

Формальное описание индексирования. Для индексации содержимого сайта создаётся специальная база данных, в которую заносится информация, собираемая поисковыми алгоритмами со страниц сайтов, а процесс добавления собранной информации в базу называется индексацией. Далее собранные данные определённым образом обрабатываются, и создается индекс–выжимка из документов. При этом поисковыми запросами учитываются текстовое наполнение, внутренние и внешние ссылки, графические и некоторые другие объекты. Когда пользователь сформирует свой запрос поисковой системе, происходит обращение к этой базе данных. В результате поисковому движку для выполнения поискового запроса не нужно обрабатывать контент каждый раз заново — поиск выполняется по индексу. Сам процесс индексации можно сформировать, выполняя следующие действия (рисунок 2.):



Рисунок 2. Описание работы механизма индексации

Пункты 1-3, показанные на рисунке 2, можно применить для получения формы из рисунка 1 пользовательского запроса для сравнения с индексированными данными.

Как сказано выше, при организации поиска с учётом языковой морфологии надо использовать два специально созданных морфологических модуля, называемые стеммером и лемматизатором.

Стемминг - это приведение слова к базовой форме с исключением из заданного слова окончаний и суффиксов, и в конечном итоге оставшаяся часть слова удовлетворяет условиям идентичности для всех грамматических форм слова. Надо иметь в виду, что не во всех языках можно применить стеммер, потому что данную технологию можно использовать, только если язык при словоизменении реализуется через аффиксы.

Обычно при организации поиска текста с имитацией учёта морфологии применяют стемминг. Здесь под термином «имитации» надо понимать появление ошибок и нерелевантных результатов, которые могут появиться, если применять только стеммер. При использовании стеммера не все слова проводятся к базовой форме. На примере можно увидеть это, определяя базовую форму заданного слова «хисса» («хисса» – базовая форма «хис», с узб. языка, вклад), по ошибке выделяются слова «хисобот», «хис-хаяжон», «хис-туйғу», «хисоб» (рисунок 3.).

Исходное предложение

Исломшунослик	илмига	салмоқли	хисса	қўшган	аждоқларимиз	асарлари
↓	↓	↓	↓	↓	↓	↓
Ислом	илм	сал	хис	қўш	аждоқ	асар

Рисунок 3. Пример использования стеммера

Для решения проблемы нужны два дополнительных модуля грамматического словаря – лемматизатор (склонения) и флексер (спряжения) (рисунки 4-5.).

Исходное предложение

Исломшунослик	илмига	салмоқли	хисса	қўшган	аждоқларимиз	асарлари
↓	↓	↓	↓	↓	↓	↓
Ислом	илм	салмоқ	хисса	қўшмоқ	аждоқ	асар
		улкан				

Рисунок 4. Простая лемматизация

Исходное предложение

Исломшунослик	илмига	салмоқли	хисса	қўшган	аждоқларимиз	асарлари
↓	↓	↓	↓	↓	↓	↓
Исломшунос	илм	салмоқ	хисса	қўшмоқ	аждоқ	асар

Рис. 5. Глубокая нормализация текста с приведением к существительным с помощью тезауруса.

Как вид из рисунков 4-5, морфологический анализатор неявно даёт на выходе информацию о леммах, так как он распознаёт для каждого исходного слова принадлежность к определённой словарной статье, так как у словарной статьи можно получить начальную форму слова, и таким образом решить задачу. Конечно, побочным эффектом такого алгоритма будет учёт контекста слова и отбрасывание альтернативных вариантов распознавания.

7. Выводы

Сформированы основные принципы работы поисковой машины с поддержкой полнотекстового поиска. Для улучшения результата поиска можно сказать, что важно правильно отделять основу слов, и этот процесс значительно облегчает процесс поиска. При правильном морфологическом анализе слов улучшается результат при поиске запроса.

С помощью лемматизатора нужно приводить слова к базовой форме, после этого надо сопоставить слова со стемом, уточняя результат с помощью лемматизации. А используя второй модуль, можно выдавать грамматические формы заданного слова на основе базовой. С помощью этих

модулей при поиске можно уточнять выдачи, проверив найденные фрагменты по набору форм ключевого слова.

Из проведённых примеров видно, что применение глубокой нормализации текста с приведением к существительным с помощью тезауруса даёт более надёжные результаты.

Литература:

1. Глаговский К.А., Моисеев Л.Е., Рудельсон Л.Е. Многокритериальный поиск данных на основе информационных образов // Научный вестник МГТУ ГА. -2013. №195.
2. Земцов А.Н., Зунг Хань Чанг. Об эффективности поиска данных в веб приложениях // Инженерный вестник Дона. – 2017. №3.
3. Andrew Aksyonoff Introduction to Search with Sphinx. USA: O’Reilly Media, Inc, 2011.
4. Ахтамова С.С. алгоритмы поиска данных // Современные наукоемкие технологии. – 2007. – № 3. – С. 11-14; URL: <http://www.top-technologies.ru/ru/article/view?id=24620> (дата обращения: 10.11.2019).
5. Gonzalo Navarro Approximate Text Searching: PhD/ Computer Science, Chile, 1998.
6. Обзор методов и алгоритмов полнотекстового поиска. // <http://www.ict.nsc.ru/ru/education/seminar/obzor-metodov-polnotekstovogo-poiska> (дата обращения: 5.11.2019).
7. Колосов А. П. Алгоритм полнотекстового поиска с обучением на основе статистических данных // Известия ТулГУ. Технические науки. 2011. №6-2. URL: <https://cyberleninka.ru/article/n/algoritm-polnotekstovogo-poiska-s-obucheniem-na-osnove-statisticheskikh-dannyh> (дата обращения: 11.12.2019).
8. Веретенников А. Б. Применение трёхкомпонентных ключей для полнотекстового поиска с учётом расстояния с гарантированным временем отклика// Вестн. ЮУрГУ. Сер. Выч. матем. информ., 2018, том 7, выпуск 1, 60–77 DOI: <https://doi.org/10.14529/cmse180105>
9. Соченков И.В., Суворов Р.Е. Сервисы полнотекстового поиска в информационно - аналитической системе // Информационные технологии и вычислительные системы. 2013. №2.
10. Равшанов Н., Шафиев Т.Р. Использование специальных алгоритмов для формирования запросов пользователей в интеллектуальных системах // Доклады республиканской научно-технической конференции “Современное состояние и перспективы применения информационных технологий в управлении”, Ташкент: 2019. С. 372-378.
11. Kayest M. and Jain S. K. “Optimization driven cluster based indexing and matching for the document retrieval,” J. King Saud Univ. - Comput. Inf. Sci., no. xxxx, pp. 0–10, 2019, DOI: 10.1016/j.jksuci.2019.02.012.
12. Moloshnikov I. A., Sboev A. G., Rybka R. B., and Gydovskikh D. V. An Algorithm of Finding Thematically Similar Documents with Creating Context-semantic Graph Based on Probabilistic-entropy Approach, vol. 66. Elsevier Masson SAS, 2015.
13. Paiva S. “A Fuzzy Algorithm for Optimizing Semantic Documental Searches,” Procedia Technol., vol. 9, pp. 1–10, 2013, DOI: 10.1016/j.protcy.2013.12.001.
14. Eto M. “Extended co-citation search: Graph-based document retrieval on a co-citation network containing citation context information,” Inf. Process. Manag., vol. 56, no. 6, p. 102046, 2019, DOI: 10.1016/j.ipm.2019.05.007.
15. Кизянов А. Ф. Рассмотрение возможности расширения индексации текста на основе стимминга с возможностью автоматической индексации текста на произвольном языке // Вестник Адыгейского государственного университета. 2007. № 4. С. 56–62.