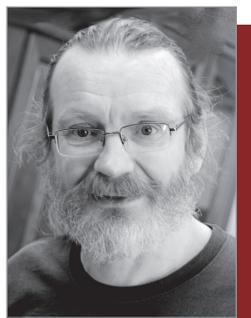


# Автоматическая предметизация короткой фразы методом сравнения с библиографическими записями в электронном каталоге

*Предлагается новый метод автоматической тематической классификации по полноразмерному естественному рубрикатору, могущий быть примененным к сверхкоротким текстам. В основе метода — использование уже принятых классификационных решений, заимствуемых из библиографических записей. Найденные решения уточняются информацией, получаемой от классификационной системы, и генерализируются различными способами. Рассмотрено четыре способа получения классификационного решения, отличающихся числом отказов и долей неверных результатов. Продемонстрирована успешность метода для предметизации заголовков произведений делениями ББК. Рассмотрены проблемы библиографического каталога и классификационной системы, ухудшающие результат. Указаны возможные отрицательные последствия использования автоматической классификации для качества библиографического каталога.*



**Владимир Юрьевич  
Калачихин,**  
начальник отдела развития  
и использования когнитивных  
технологий  
Российской государственной  
библиотеки

**Ключевые слова:** предметизация, систематизация, ББК, Библиотечно-библиографическая классификация, распознавание смысла, сверхкороткие тексты.

## Введение

Вместе с ростом объемов коллекций текстов в электронной форме вне библиотечной среды появилась необходимость какого-то упорядочивания и организации доступа к этим текстам, а вместе с тем — и понимание высокой трудоемкости данного процесса. Имеющийся обширный опыт использования поисковых систем в Интернете

принес также понимание недостаточности простого контекстного поиска. Поэтому интерес многих специалистов оказался направленным на задачи автоматической классификации и систематизации документов. Если судить по библиотеке диссертаций Российской государственной библиотеки (РГБ), то бурный расцвет исследований в этой области пришелся на 2000—2005 годы. Однако столь же быстро произошел и спад, и уже к 2010 г. диссертаций на темы автоматической систематизации текстов почти нет. К настоящему времени очевидно, что вся эта активность не оставила никакого следа в реальной библиотечной деятельности ни в России, ни за рубежом.

Основные развиваемые в начале 2000-х гг. идеи: искусственный рубрикатор (типичная работа, например, [6]), классификация по естественному рубрикатору [3] с использованием статистических методов или методов кластеризации — давали неоднозначные или прямо непригодные результаты даже в очень упрощенных условиях. Искусственный рубрикатор оказывался несоотносим с традиционным пониманием рубрикатора и поэтому неприменим на практике [6], классификация по естественному рубрикатору получалась чрезмерно упрощенной (использовалось 600 рубрик, тогда как Таблицы ББК для массовых библиотек (самый простой вариант) содержат 2,5 тыс. рубрик) [3].

Постепенно сложилось понимание, что содержательного результата можно ожидать только от очень простых методов, опирающихся на практическое использование языка. Так, в работе «Алгоритм и программа индексации текста по естественному рубрикатору» — [4] автор предпринял попытку автоматической систематизации электронного текста по полноразмерному рубрикатору (часть Полных таблиц ББК, около 60 тыс. рубрик), исходя из естественных свойств текста. Получены обнадеживающие результаты.

Другой подход, также основанный на естественных свойствах текста, применен в технологии «Семантическое зеркало» [подробнее см. 1; 5; 7]. Утверждается, что эта технология уже применяется в целом ряде реально существующих интернет-проектов. Суть технологии в соотношении короткой фразы «как есть» с набором слов и словосочетаний, приписанных рубрикам естественного рубрикатора. Рубрикатор состоит примерно из 3 тыс. рубрик.

Авторами технологии отдельно акцентируется внимание на способности «Семантического зеркала» классифицировать сверхкороткие тексты — отдельные фразы и предложения. На самом деле, удачно классифицируются не любые короткие фразы, а только заголовки и поисковые запросы: «именно краткость таких текстов, как заголовки и поисковые запросы, практически исключает случайное попадание в них тематически маркированных слов и особенно словосочетаний» [1].

На этом обстоятельстве и основаны предлагаемые нами алгоритм и программное обеспечение, способные классифицировать заглавия произведений по полноразмерному естественному рубрикатору примерно по тому же принципу, что и «Семантическое зеркало», но совершенно другим способом.

## Описание алгоритма

### Идея

В поиске способов автоматической систематизации (или предметизации) нами была исследована идея систематизации «по аналогии» с существующими решениями по систематизации. Суть идеи в том, что имея заглавие произведения, не снабженного индексом ББК, можно искать в каталоге библиографическую запись на произведение с похожим заглавием, имеющую индекс ББК, и считать, что данный индекс будет достаточно подходящим и для исходного произведения.

### Воплощение

Достаточно очевидно, что искать только похожие заглавия — бесперспективно. Заглавия произведений довольно уникальны и коротки, тогда как библиографическая запись целиком включает и ключевые слова, и аннотацию. Поэтому разумно искать слова из заглавия систематизируемого

произведения среди всех полей библиографической записи.

Разумеется, как обычно бывает при простом текстовом поиске, наибольший вклад в совпадение дают слова общей лексики, поэтому для улучшения результатов сравнения было предпринято следующее:

- повышение содержательности запроса;
- отбор только тех ответов из каталога, индекс ББК которых совпадал бы с индексом ББК, полученным из аналогичного запроса к таблицам ББК.

Второе направление выглядело наиболее перспективным.

Повышение содержательности запроса осуществляется удалением из него слов общей лексики. Для этого используется имеющийся в нашем распоряжении словарь, содержащий около тысячи стоп-слов. Данный словарь применяется нами в других проектах и его адекватность проверена.

Отбор ответов осуществляется пересечением множеств индексов ББК, полученных из ответа каталога, с множеством индексов ББК, полученных из ответа информационной системы

ПИКВИКС (<http://lbc.rsl.ru/>), содержащей несколько вариантов таблиц ББК. Запрашиваются все имеющиеся в нашем распоряжении варианты ББК: массовые таблицы, часть полных и средних таблиц, а также так называемые разделители ГСК (Генерального систематического каталога РГБ) — иерархия индексов ББК, используемая для построения систематического каталога.

Существенным является вопрос о том, какое количество ответов поисковой системы брать в обработку. Если взять мало — высока вероятность того, что наиболее подходящие индексы ББК будут упущены. Если взять много — возрастает вероятность случайных совпадений и неприемлемо возрастет время работы.

Для информационной системы ПИКВИКС с таблицами ББК мы ориентировались на релевантность ответов. Известно, что при любом разумном определении понятия релевантности резуль-

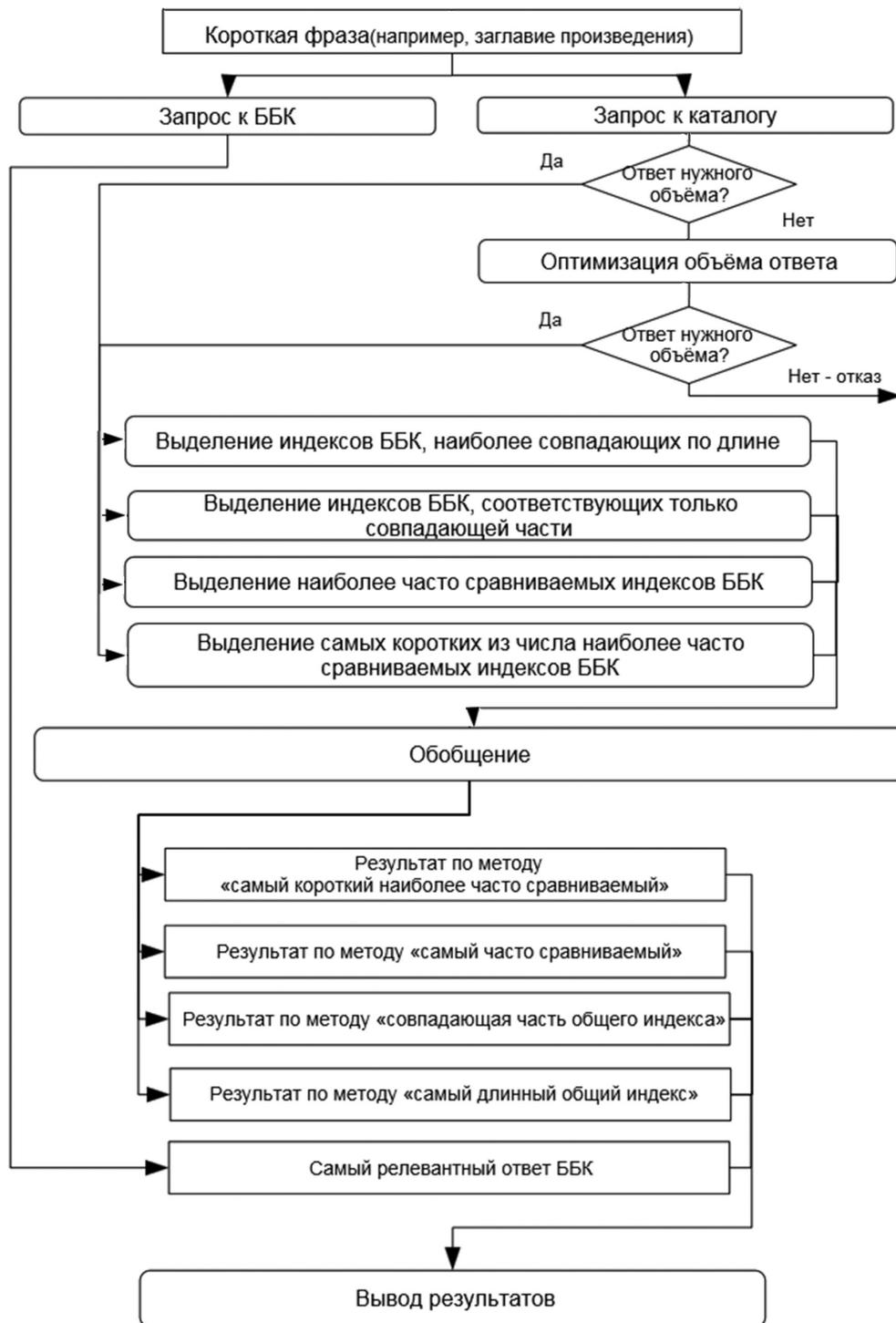


Рис. 1. Алгоритм сопоставления короткой фразы с делениями ББК

татов полнотекстового поиска эта величина в ответах сначала быстро падает, а потом — медленно уменьшается (эмпирический «закон Парето — Ципфа — Кудрина», имеющий термодинамическое обоснование [2]). Иными словами, ответов с высокой релевантностью мало, а с низкой — много. Это, в основном, общие слова и случайные совпадения. Поэтому естественно брать для анализа ответы с релевантностью от максимальной до конца быстрого падения. Однако анализ характера падения релевантности для каждого ответа был бы слишком трудоемок, поэтому за указанную границу была принята экспериментально определенная константа.

К сожалению, для ответов каталога такой подход неприемлем, так как используемый доступ к электронному каталогу не всегда позволяет получить результат, отсортированный по степени релевантности запросу. Поэтому пришлось принять меры по составлению запроса таким образом, чтобы ответ был достаточен по объему, но не слишком велик для получения его от электронного каталога целиком. Если же это оказывалось невозможным, то из ответов каталога для анализа отбиралось фиксированное число первых ответов, определенное экспериментально. К сожалению, данный подход дает гораздо худшие результаты, чем отбор ответов с релевантностью не хуже заданной.

Естественно, индексы ББК из библиографического каталога и собственно из таблиц почти никогда не совпадают полностью, поэтому в итоговое множество попадают индексы, отобранные по двум критериям:

- индекс ББК из таблиц, наиболее совпадающий по длине с индексом ББК из каталога (метод «самый длинный общий индекс»);
- индекс ББК из таблиц, наиболее часто сравниваемый с индексом ББК из каталога (метод «самый часто сравниваемый»).

Обнаружилось, что по обоим критериям часто выделяется более одного индекса, причем некоторые из выделенных индексов являются прямыми потомками одного вышележащего деления, а другие сами являются вышележащими, хотя и не непосредственными предками выделенных. Поэтому было сочтено целесообразным обобщать результат, заменяя несколько нижележащих делений одним их непосредственным предком, а другие нижележащие просто игнорировать, если в ответе присутствует вышележащее деление.

Кроме того, на основе каждого из множеств индексов ББК, отобранных по двум критериям, создавались дополнительные множества:

- на основе отобранных по методу «самый длинный общий индекс» — индексы ББК, не превышающие по длине общую для ответов ББК и ответов каталога часть индекса (метод «совпадающая часть общего индекса»);
- на основе отобранных по методу «самый часто сравниваемый» — индексы ББК, самые короткие из них (метод «самый короткий наиболее часто сравниваемый»).

В последнем случае предполагается, что более короткий — более общий по смыслу индекс. Это не обязательно так, но другого способа разделить деления на выше- и нижележащие в своих ветвях рубрикатора — в ББК нет.

Таким образом, нами предложен алгоритм (см. рис. 1), позволяющий любой короткой фразе двумя способами и двумя разновидностями способов сопоставить одно или несколько соответствующих этой фразе по смыслу делений ББК. Алгоритм позволяет идентифицировать ситуацию «отказа» — когда деления не могут быть сопоставлены по тем или иным причинам.

Алгоритм реализован в виде программного продукта, с работой которого можно ознакомиться по адресу <http://lbc.rsl.ru/anysy/>

## Исследование работы алгоритма

### Цель

Целью исследования было выяснить, дает ли предлагаемый алгоритм результаты лучше, чем простой текстовый поиск по таблицам ББК, а также какой из методов выделения понятий и их вариантов лучше.

### Методика

Из корпуса текстов случайным образом выбирался файл с текстом на русском языке. Далее по заглавию произведения уточнялось, что тема произведения соответствует имеющимся в нашем распоряжении разделам ББК. После чего заглавие произведения систематизировалось программой.

Результат систематизации оценивался по качественной шкале путем просмотра текста произведения, его оглавления, аннотации и т. д. примерно в соответствии с тем, как это делают систематизаторы. При этом оценивалось соответствие результатов работы программы именно смыслу произведения, а не смыслу его заглавия. Разумеется, соответствие смысла заглавия смыслу произведения решительно влияло на результат работы.

При анализе качества работы программного комплекса использовалась следующая шкала:

- 1 балл — результат полностью неправильный;
- 2 балла — результат неточный или слишком общий, или слишком частный, или не совсем правильный, или относящийся к побочным темам или аспектам документа;
- 3 балла — результат отражает главную тему документа слишком обще или не очень точно, но может быть использован для тематического поиска в двух-трехуровневых рубриках;
- 4 балла — результат отражает главную тему документа с точностью, достаточной для тематического поиска;
- 5 баллов — результат полностью и адекватно отражает главную тему документа.

Отсутствие в результате индексов из вспомогательных таблиц (например, «учебное пособие») не являлось причиной снижения оценки.

При оценке игнорировались некоторые проблемы ББК, как то архаичность структуры и формулировок или неполнота разделов. Например, результат «3.9.7.3.2.3.3-01.8 Программирование для информационных машин» признавался полностью правильным (5 баллов) для книги по программированию, несмотря на то что это деление является подразделом деления «Специализированные и управляющие электронные вычислительные машины дискретного действия», а не деления «Универсальные электронные вычислительные машины дискретного действия», как было бы правильно, потому что деления «Программирование для информационных машин» внутри раздела «Универсальные электронные вычислительные машины дискретного действия» не существует.

Аналогично результат со смыслом «методика преподавания английского языка» для соответствующего методического пособия признавался адекватным, хотя это деление было подразделом методики школьного преподавания, а книга относилась к высшей школе. Но в разделе ББК о высшей школе нет деления о преподавании языков. Поэтому результат получил 5 баллов, а не 2, как было бы при более правильном построении классификационной системы. Такой подход тем более оправдан, что многие разделы ББК интенсивно перерабатываются, и у нас нет актуальных их вариантов.

Одновременно с результатами работы алгоритма фиксировался и результат простого текстового поиска по таблицам ББК — самый релевантный ответ информационной системы ПИКВИК. С этим результатом и сравнивался результат работы алгоритма.

Оценки работы по четырем методам и самый релевантный ответ ББК осреднялись. Ситуация отказа баллами не оценивалась и в вычислении среднего не участвовала.

Достоверность различий между вычисленными средними оценивалась по t-критерию Студента. Применимость t-критерия не оценивалась, поэтому он использовался только для качественной оценки достоверности.

Источником библиографических записей являлся электронный каталог РГБ, доступ к которому осуществлялся по протоколу Z39.50.

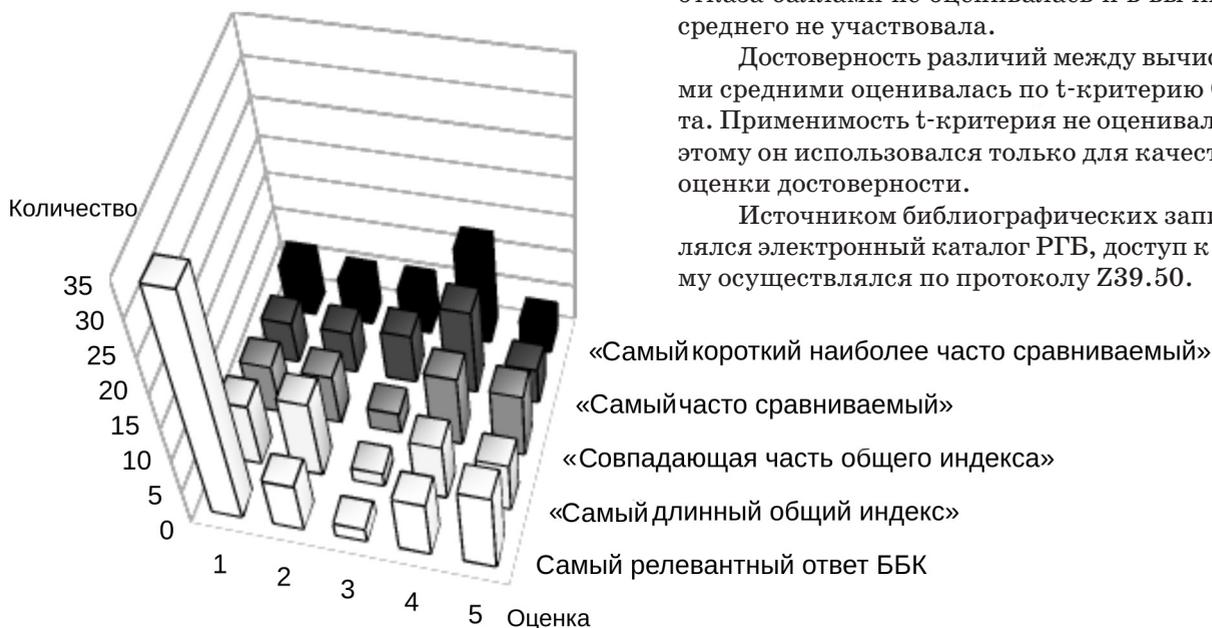


Рис. 2. Распределение оценок

### Условия

Работа алгоритма была исследована на 61 заглавии произведений из состава LibraryGenesis, в основном на техническую и научную тематику, случайным образом отобранных из электронного каталога. В библиотеке LibraryGenesis насчитывается более 200 тыс. наименований на русском языке по широкому спектру тем, поэтому в выборку попали произведения различной тематики. В нашем распоряжении были тексты этих произведений, так что имелась возможность оценить степень соответствия выделенных алгоритмом понятий ББК реальному содержанию текстов. Степень соответствия оценивалась по шкале таким образом, что результат с оценкой менее 3 баллов не является практически пригодным, а с оценкой 3 и более — может быть использован для реальной систематизации с теми или иными оговорками.

Поскольку оценивание производилось с точки зрения качества систематизации, соответствие смысла заглавия смыслу произведения вынужденно влияло на оценку.

Вместе с результатами работы алгоритма фиксировался и аналогично оценивался самый релевантный ответ системы ПИКВИКС с таблицами ББК.

Ожидалось, что простой текстовый поиск по ББК даст среднюю оценку результата менее 3 баллов, а по крайней мере один из методов работы предлагаемого алгоритма — больше 3 баллов.

### Результаты

#### Распределение оценок

Распределение количества оценок, полученных результатами работы каждого из методов и простого текстового поиска в ББК, приведено на рис. 2. Видно, что алгоритм довольно эффективно обрабатывает ситуацию отказа, когда приемлемые понятия ББК не могут быть найдены — число оценок в 1 балл среди оценок работы алгоритма значительно меньше, чем среди оценок текстового поиска. Также видно, что все методы, кроме метода «самый длинный общий индекс» показывают преобладание содержательных оценок в 4 и, отчасти, 5 баллов. Все методы показывают и определенную

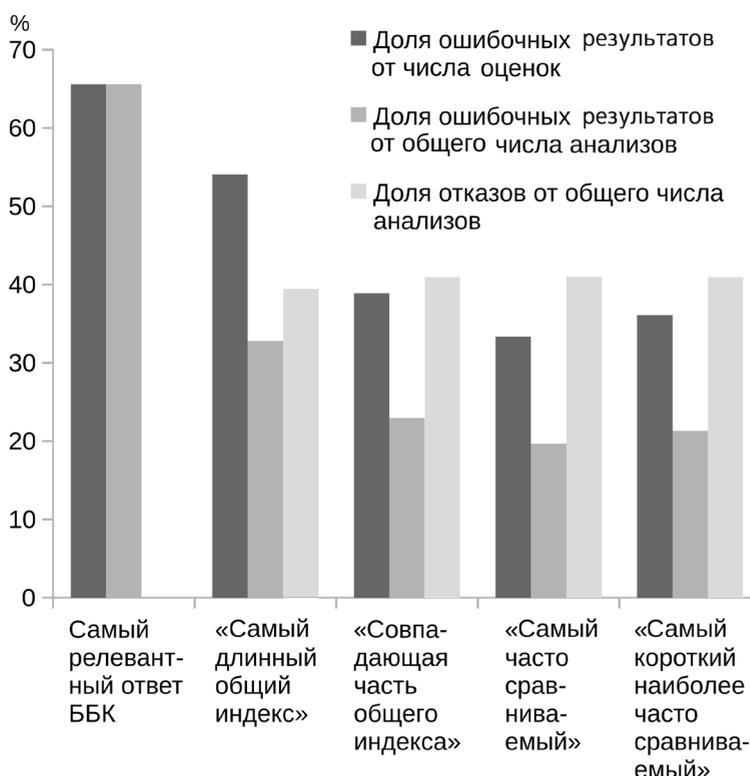


Рис. 3. Ошибки и отказы по выбранным методам

избирательность — они действительно выделяют понятия более содержательно, чем простой текстовый поиск. Однако общее распределение оценок для методов выбора по длине индекса похоже на таковое для простого текстового поиска, что говорит о том, что избирательность этих методов может быть обусловлена в основном реакцией на ситуацию отказа.

Напротив, методы по числу сравнений показывают другой характер распределения оценок. Это заставляет предполагать, что избирательность этих методов обусловлена, кроме реакции на отказ, еще и работой самого метода.

#### Эффективность и достоверность

Средние оценки по всем результатам анализа следующие:

- самый релевантный ответ ББК — 2,3 балла;
- метод «самый длинный общий индекс» — 2,8 балла;
- метод «совпадающая часть общего индекса» — 3,2 балла
- метод «самый часто сравниваемый» — 3,1 балла;
- метод «самый короткий наиболее часто сравниваемый» — 3,0 балла.

Видно, что простой текстовый поиск действительно дает результат в среднем менее 3 баллов, а три метода работы алгоритма — более или равный 3 баллам.

Исследование достоверности различия этих средних с помощью *t*-критерия Студента показало, что значения среднего для ответов ББК достоверно отличаются от всех остальных значений, значения по методу «самый длинный общий индекс» достоверно отличаются от остальных, кроме результатов по методу «самый короткий наиболее часто сравниваемый», остальные средние достоверно не различаются.

Таким образом, можно утверждать, что работа алгоритма дает статистически достоверно лучший результат, чем простой текстовый поиск. При этом результат простого текстового поиска в среднем хуже, чем приемлемо для практического использования, а результат работы алгоритма —

в среднем достаточен (в соответствии с принятой шкалой).

#### Ошибки и отказы

Доля неверных результатов и число отказов для различных методов приведены на рис. 3. Видно, что доля отказов примерно одинакова для всех вариантов работы алгоритма и составляет 40% от общего числа анализов, а вот доля неверных оценок минимальна у метода «самый часто сравниваемый», и составляет 33%, что вдвое меньше, чем у простого текстового поиска (65%), если считать от числа оценок (без отказов). Если же считать от общего числа анализов, включая и отказы, доля неверных результатов для этого метода составляет 20% против 65% простого текстового поиска.

Для подсчета общей доли правильных результатов к этому числу надо прибавить 40% отказов. В результате простой текстовый поиск дает примерно 35% правильных результатов, а лучший вариант нашего алгоритма — 40%. Однако для целей автоматической систематизации интересна доля наименований, которая может быть автоматически проиндексирована, и с этой точки зрения отказ — лучше, чем неверный результат. Поэтому правильность работы следует считать по числу неверных результатов, и здесь наш алгоритм показывает решительное преимущество перед простым текстовым поиском — 20% неверных результатов против 65%.

#### Варианты ББК

Наиболее полный вариант ББК, имеющийся в нашем распоряжении — это так называемые разделители ГСК — плоский рубрикатор, построенный из карточек-разделителей Генерального систематического каталога РГБ. Деления этого рубрикатора — это составные индексы ББК с их словесной формулировкой. Никаких дополнительных сведений для этих делений нет.

Кроме разделителей ГСК использовались Полные таблицы ББК в объеме примерно половины основного ряда (со вспомогательными табли-

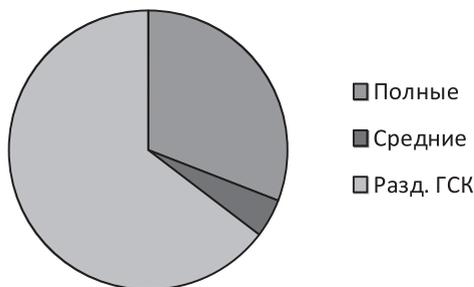


Рис. 4. Абсолютная частота использования разных вариантов Таблиц ББК

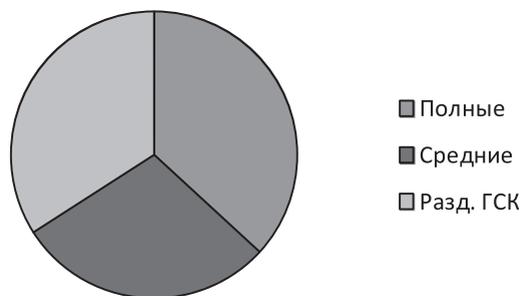


Рис. 5. Относительная частота использования разных вариантов Таблиц ББК

цами), Средние таблицы в объеме примерно трети основного ряда (также со вспомогательными таблицами) и Массовые таблицы ББК полностью. Все эти варианты в составе деления, кроме словесной формулировки, содержат методические указания и, иногда ключевые слова.

Поскольку разделители ГСК охватывают наиболее полный объем понятий, деления этого рубрикатора наиболее часто встречались в результатах работы алгоритма (рис. 4). Однако оказалось, что относительно имеющегося объема рубрикаторов деления из разных рубрикаторов встречаются одинаково часто (рис. 5). Очевидно, что наибольший вклад в релевантность поиска по рубрикатору вносит словесная формулировка деления, а наличие методических указаний и ключевых слов не очень существенно. Более того, анализ некоторых неудачных результатов работы алгоритма показал, что имеет место некорректное указание ключевых слов в Полных таблицах ББК: слова, относящиеся к вышележащим делениям, были указаны и для нижележащих. Это заметно ухудшало общую избирательность работы алгоритма.

#### **Скорость**

Анализ одной короткой фразы занимает порядка секунд на машине начального серверного класса. Наибольшие задержки вызывает ожидание ответа электронного каталога.

#### **Проблемы**

##### **Качество каталога и таблиц ББК**

Анализ неверных результатов работы алгоритма показывает, что причиной может быть нерелевантный ответ ББК, нерелевантный ответ каталога или и то и другое вместе. Основная (по числу неверных результатов) причина нерелевантного ответа системы ПИКВИКС с таблицами ББК — это неверная словесная формулировка деления ББК в разделителях ГСК. Достаточно часто в качестве словесной формулировки деления применяется целая цепочка формулировок, начиная от какого-нибудь общего деления или даже от корня рубрикатора. Например, формулировка деления Ш5(2=Р)75-646.41Ар выглядит так:

*Филологические науки. Художественная литература — Художественная литература. Литературоведение — Мировая литература народов отдельных стран — Литература СССР — Литература отдельных народов СССР — Русская литература — Русская советская литература — Литература с середины 1950-х гг. — Произведения художественной литературы — Художественная проза — Очерки — Сборники произведений разных авторов — Путевые очерки и очерки об отдельных странах и местностях — Очерки о родной стране — Арктика*

Очевидно, что это деление будет найдено и по слову «литература», и по слову «очерки», и по слову «проза». А также — Арктика, СССР и 1950.

Аналогичная проблема цепочек имеется и в каталоге. Непонятно с какой целью значительное число библиографических записей вместе с индексами ББК имеют и полную цепочку словесных формулировок делений ББК, начиная от корня. Например, библиографическая запись на диссертацию «Конвективно-плёночное охлаждение в сверхзвуковом высокотемпературном потоке, численное решение сопряженной задачи» содержит два индекса ББК: 0651-046,0 и 0551.41-046,0 и две цепочки:

*Транспорт — Межпланетные полеты — Космонавтика — Ракетные двигатели космических летательных аппаратов — Системы охлаждения*

*Транспорт — Воздушный транспорт — Авиация — Авиационные двигатели — Авиационные воздушно-реактивные двигатели — Авиационные газотурбинные двигатели — Системы охлаждения*

— т. е. 10 высокосодежательных терминов, относящихся к совершенно разным областям техники, и, к тому же не имеющих прямого

отношения к теме диссертации! Очевидно, такая библиографическая запись не может быть корректно использована в работе нашего алгоритма.

#### Доступ к каталогу

Серверы Z39.50 имеют каждый свои особенности, поскольку в основном создавались для выполнения узких функций, а задача полноценной поддержки протокола перед разработчиками, как правило, не стояла. Поэтому разные серверы имеют различные отхождения от спецификации протокола и различные особенности его реализации. В частности, весьма существенной особенностью является то, что одни серверы возвращают отсортированный по релевантности результат, другие — не отсортированный. К сожалению, с одними и теми же настройками программа будет давать разный результат при работе с разными источниками каталога. По возможности мы постарались подобрать наиболее общие параметры (пусть и в ущерб скорости работы), но тем не менее при смене сервера Z39.50 требуется отдельно оценить адекватность параметров именно для этого сервера.

#### Деграция каталога

Очевидно, что работа алгоритма дает результат систематизации не лучше, чем уже существует в каталоге. Поскольку в работе алгоритма используется обобщение результатов, то в среднем результатами работы алгоритма являются более общие деления ББК, чем в среднем по каталогу.

Таким образом, если возвращать результаты работы алгоритма в каталог, а потом снова использовать, можно ожидать постепенного снижения качества систематизации в среднем по каталогу. Непонятно на сегодня, будет ли этот процесс непрерывным и приведет со временем к полной деграции каталога или стабилизируется на каком-то уровне общности. Мы склоняемся ко второму варианту, но эта проблема требует исследования.

#### Перспективы

Предлагаемый алгоритм, реализованный в виде программного обеспечения, позволяет, в принципе, осуществлять автоматическую систематизацию нехудожественных произведений по их заглавию. Но для корректной работы алгоритма требуется наличие качественно выполненных таблиц ББК в полном объеме, и не имеющих лишней информации библиографических записей с надлежащего качества индексами ББК. При

соблюдении этих условий возможна полностью автоматическая систематизация по основной теме не менее двух третей наименований, с ошибочной систематизацией — менее пятой части наименований. При систематизации диссертаций можно ожидать лучшего результата.

Если вместо одной короткой фразы анализировать сжатое изложение содержания документа — аннотацию или (автоматически составленный) реферат — то, вероятно, результат будет еще лучше.

#### Список источников

1. Белов А.А. Автоматическое распознавание тематики сверхкоротких текстов [Электронный ресурс] / А.А. Белов, М.М. Волович. — URL: <http://www.dialog-21.ru/dialog2007/materials/html/05.htm>
2. Бялко А.В. Эмпирический «закон Парето—Ципфа—Кудрина» и общая теория конкуренции [Электронный ресурс] / А.В. Бялко, Б.А. Трубников, О.Б. Трубникова // Общая и прикладная ценология. — 2007. — № 4. — С. 20—24. — URL: <http://www.kudrinbi.ru/public/20151/index.htm>
3. Дунаев Е.В. Автоматическая рубрикация веб-страниц в интернет-каталоге с иерархической структурой [Электронный ресурс] / Е.В. Дунаев, А.А. Шелестов. — URL: <http://hdl.handle.net/10995/1419>
4. Калачихин В.Ю. Алгоритм и программа индексации текста по естественному рубрикатору [Электронный ресурс] / В.Ю. Калачихин. — URL: <http://lbc.rsl.ru/aind/autoindex.html>
5. Маркович Я. Практическое применение семантического анализа для фильтрации трафика [Электронный ресурс] / Я. Маркович, М. Волович. — URL: <http://profyclub.ru/docs/249>
6. Пескова О.В. Автоматическое формирование рубрикатора полнотекстовых документов [Электронный ресурс] / О.В. Пескова. — URL: [http://peskova.ru/Papers/opeskova\(RCDL2008\).pdf?utm\\_source=searchengine&utm\\_medium=organic&utm\\_campaign=pdf](http://peskova.ru/Papers/opeskova(RCDL2008).pdf?utm_source=searchengine&utm_medium=organic&utm_campaign=pdf)
7. Технология «Семантическое зеркало» [Электронный ресурс]. — URL: <http://sm.ashmanov.com/product/#sm>

Контактные данные:

119019, Москва, ул. Воздвиженка, д. 3/5;  
e-mail: [KalachikhinVY@rsl.ru](mailto:KalachikhinVY@rsl.ru)