

АВТОМАТИЧЕСКОЕ ИСПРАВЛЕНИЕ ОПЕЧАТОК В ПОИСКОВЫХ ЗАПРОСАХ БЕЗ УЧЕТА КОНТЕКСТА

Панина М. Ф. (mar-fed@yandex-team.ru),
Байтин А. В. (baytin@yandex-team.ru),
Галинская И. Е. (galinskaya@yandex-team.ru)

Яндекс, Москва, Россия

Анализируя ошибки в поисковых запросах нетрудно заметить, что большая часть из них имеет однозначное исправление, не зависящее от словарного окружения, и может быть исправлена в автоматическом режиме. В данной работе мы попытались выделить классы ошибок, которые можно исправлять автоматически, определить долю контекстно-независимых исправлений, и для выбранного множества ошибок разработать классификатор, позволяющий разделить исправления на надежные (пригодные для автоматической замены) и малонадежные (пригодные только для подсказки).

В качестве кандидатов для исправлений были использованы подсказки поискового спелл-чекера, знакомые пользователям поисковых систем по сообщению «Возможно, вы имели в виду...». Для обучения классификатора были использованы лексические и статистические признаки словарного (бесконтекстного) уровня.

Проведенные эксперименты показали высокую эффективность признаков и возможность настройки классификатора на заданный уровень точности. Применение предложенного метода автокоррекции тривиальных опечаток может значительно повысить качество исправления ошибок в поисковых запросах.

Ключевые слова: исправление опечаток в поисковых запросах, автоисправление опечаток, машинное обучение, оценка надежности, контекстно-независимые ошибки

CONTEXT-INDEPENDENT AUTOCORRECTION OF QUERY SPELLING ERRORS

Panina M. F. (mar-fed@yandex-team.ru),
Baytin A. V. (baytin@yandex-team.ru),
Galinskaya I. E. (galinskaya@yandex-team.ru)

Yandex, Moscow, Russia

While analyzing errors in the search queries, it is easy to notice that the most part of query spelling errors are trivial typos. Such errors usually do not depend on the surrounding words and their correction can be done in the automatic mode. In this work we tried to define a class of query spelling errors that can be corrected automatically. For the selected class we developed a classifier dividing corrections into reliable (suitable for automatic query spelling correction) and low-reliable (suitable only for the query spelling suggestion). As candidates for autocorrections we used query speller suggestions familiar to the users of search engines by «Did you mean...» function. For the classifier training we used typical lexical and statistical features. The experiments showed high performance of the word-level features and the ability to configure the classifier for a given level of accuracy. The application of the proposed method of trivial typo correction can significantly improve the quality of the query spelling errors correction.

Key words: query spelling correction, autocorrection, machine learning, confidence estimation, spellchecker

1. Введение

Анализ логов современных поисковых систем показывает, что пользователи делают ошибки в 10–15% поисковых запросов [8,5]. Поскольку поиск по искаженному запросу обычно приводит к нерелевантной выдаче и негативно влияет на общее качество поиска, поисковые системы уделяют большое внимание проблеме исправления опечаток в запросе и стараются исправлять их на всех этапах поискового процесса — начиная с корректировки запроса во время набора и заканчивая переформулировками во время показа результатов поиска.

Все способы исправления запросов в той или иной форме используют поисковый спелл-чекер — программу, которая находит в запросе ошибки и предлагает для них исправления. В данной работе мы рассмотрим две наиболее распространенные функции коррекции запросов — подсказку и автозамену.

Подсказка (Рис.1) представляет собой сообщение («*Быть может, вы искали: <...>*») и ссылку («*качок*»), по которой можно перейти на страницу с поисковой выдачей по исправленному запросу. Подсказка полезна в тех случаях, когда оригинальное написание (*качек*) является допустимым, но в то же время совпадает с опечаткой другого, более употребимого в данном контексте слова (*качок*); или когда исправление неоднозначно (для *альтар вики* исправлением может быть и *алтарь вики*, и *альтаир вики*). Подсказка работает в интерактивном режиме и является консервативным способом коррекции запроса (показывается вместе с выдачей по оригинальному запросу и не искажает результатов поиска). Недостатком подсказки является необходимость совершать дополнительные действия (кликать по ссылке). Кроме того, пользователи часто не замечают подсказку или не доверяют ей, упуская возможность увидеть более релевантную поисковую выдачу.

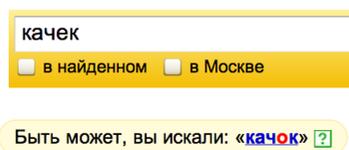


Рис. 1: Пример *подсказки*

В случае автозамены (Рис.2) пользователю показывается поисковая выдача сразу по исправленному запросу. Под поисковой строкой помещается сообщение об автоисправлении («В запросе <...> была исправлена опечатка»). Функция работает в автоматическом режиме и экономит время, которое приходилось бы тратить на ручное исправление ошибки или клик на подсказку (см. выше). На случай неверного автоисправления предусмотрена ссылка на страницу с результатами поиска по оригинальному написанию («диалог 2013»). Следует заметить, что каждое неверное автоматическое изменение запроса сильно раздражает пользователя [15], поэтому требование к точности автозамен намного выше, чем к точности подсказок.

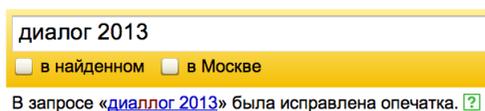


Рис. 2: Пример *автозамены*

Обнаружив в запросе ошибку, поисковая система оказывается перед выбором — какую из функций коррекции использовать. Обе функции оптимизируют полноту исправлений, однако пороги точности у подсказок и автозамен существенно различаются. Поэтому проблему выбора обычно решает не спелл-чекер, а специальный модуль (классификатор надежности), разделяющий исправления на надежные (автозамены) и ненадежные (подсказки).

Анализируя ошибки в запросах, можно заметить, что в основном они являются тривиальными опечатками и имеют очевидное исправление, не зависящее от словарного окружения. Это наблюдение приводит к идее об исправлении тривиальных ошибок в автоматическом режиме без учета информации о соседних словах. Цель данной работы состоит в оценке возможностей и разработке метода, позволяющего повысить эффективность автоматического исправления для этого класса словарных ошибок.

На фоне большого количества работ, посвященных исправлению ошибок, работ по автоисправлению относительно немного [2, 3, 9, 14]. Наиболее интересной является работа [14], посвященная построению системы для автокоррекции обычных текстов. Автоматическое исправление ошибок работает по трехуровневой модели: 1) для каждого слова спелл-чекер генерирует отранжированный список подсказок; 2) классификатор надежности

подсказки определяет является ли слово ошибкой; 3) классификатор надежности автоисправлений решает, можно ли исправить ошибку автоматически. Мы применили аналогичный подход и большую часть признаков, использованных в [14], однако надо отметить, что исправление ошибок в запросах и текстах имеют ряд существенных отличий: а) частота опечаток и степень искажения запросов гораздо выше; б) значительную долю слов в запросах составляют новые слова с неустоявшимся написанием; в) запросы в среднем состоят из 2–3 слов, поэтому для их исправления невозможно использовать широкий контекст.

Несмотря на то, что все современные поисковые системы используют функцию автозамены, нам не удалось найти ни одной статьи, посвященной непосредственно автоматическому исправлению поисковых запросов. В то же время, существуют интересные работы в области машинного перевода, где есть похожая на автозамену задача оценки качества перевода. Для того, чтобы определить степень готовности машинного перевода к ручному постредактированию или к публикации, необходимо автоматически оценивать качество перевода, имея в наличии только предложение и его машинный перевод. В работах [12,13] для решения этой задачи применяются методы машинного обучения, используются разнообразные лексические и статистические признаки. В ранних работах [12] использовался бинарный классификатор, в дальнейшем авторы перешли на пятибалльную шкалу оценки качества [13]. Важно отметить, что задача оценки качества решается независимо от задачи машинного перевода.

В данной работе для решения задачи разделения исправлений на «надежные» (автозамены) и «ненадежные» (подсказки) мы используем бинарный классификатор, подбираем для него словарные (бесконтекстные) признаки и проводим эксперимент на случайной выборке запросов.

2. Ошибки в запросах пользователей

Прежде чем приступить к разработке метода оценки надежности исправлений, необходимо произвести типизацию ошибок и их исправлений в поисковых запросах [1, 7, 8]. Поскольку ошибки каждого типа обладают особенностями и исправляются с неодинаковым качеством, для классификации надежности каждого типа нужен свой набор признаков и свои настройки классификатора.

2.1. Данные для исследования

Из дневного лога поисковой системы Яндекс случайным образом были выбраны 10 000 запросов. Из них 9242 оказались уникальными. Двум ассессорам была поставлена задача найти среди них запросы с ошибками, предложить вариант исправления и определить язык запроса. После

завершения разметки каждому ошибочному запросу Q и его исправлению C автоматически, с помощью алгоритма выравнивания Левенштейна [11], был сопоставлен набор пар $\{q_i \rightarrow c_i\}$, $i = 1...m$, где $q_i \in Q$, $c_i \in C$ — слова запроса и исправления, m — количество опечаток в запросе. Например, для пары запрос-исправление *rfr сделатьсуши дома → как сделать суши дома* количество опечаток $m=2$, массив пар слово-исправление: $\{rfr \rightarrow \text{как}; \text{сделатьсуши} \rightarrow \text{сделать суши}\}$.

Внаборе было найдено 1132 запроса с ошибками (12,2%). Распределения запросов по количеству ошибок и по языку запросов представлены в табл. 1,2:

Таблица 1. Распределение количества ошибок в запросах с ошибками

Кол-во ошибок	%
1	83,6
2	11,7
>2	4,8

Таблица 2. Распределение запросов по языкам

Язык запроса	%
Русский	80,8
Английский	7,2
Русский + Английский	8,2
другие	3,8

Из Таблиц 1,2 видно, что большая часть запросов с ошибками (83,6%) содержит всего одну ошибку и почти все запросы задаются на русском или английском языках (96,2%).

2.2. Типы ошибок

Ошибки, встречающиеся в поисковых запросах, можно отнести к одному из четырех классов:

- **Ошибки в отдельных словах** — пропуск, вставка, перестановка и замена букв в словах (напр. *сборик* → *сборник*, *статья* → *статья*, *компьютер* → *компьютер*, *модерн* → *модерн*)
- **Ошибки слитно-раздельного написания** — пропуск или вставка пробела между словами (напр. *томхарди* → *том харди*, *такси* → *такси*)
- **Неверная раскладка клавиатуры** — использование английской раскладки для набора русских слов или наоборот (напр. *сфеы* → *cats*, *rjns* → *коты*)
- **Транслитерация** — использование латинского алфавита для набора русских слов или наоборот (напр. *uragan* → *ураган*, *контракшн* → *construction*)

Стоит заметить, что, хотя неверная раскладка и транслитерация формально ошибками правописания не являются, мы относим их к искаженным написаниям, негативно влияющим на качество поисковой выдачи.

Таблица 3. Типы ошибок в запросах

Тип ошибки	%
Ошибки в отдельных словах	63,7
Ошибки слитно-раздельного написания	16,9
Неверная раскладка клавиатуры	9,7
Транслитерация	1,3
Смешанные ошибки	8,3

В данной работе мы будем рассматривать самый большой класс ошибок (63,7%) — «ошибки в отдельных словах». Для простоты изложения в дальнейшем будем называть такие ошибки словарными или просто опечатками.

2.3. Зависимость исправлений от контекста

Для того чтобы проверить предположение о том, что большая часть словарных ошибок в запросах являются тривиальными, т.е. имеющими однозначное исправление независимо от контекста, мы попытались оценить влияние словарного окружения на исправление ошибок и провели следующий эксперимент. Из набора (гл. 2.1) взяли все запросы со словарными ошибками (620 запросов), выделили из них все пары слово-исправление (754 пары), отфильтровали пары-орфоварианты¹ и пары-словоформы², из оставшихся 714 пар взяли слова и поручили аналитику исправить в них ошибки, не глядя на контекст этих слов (текст запроса). Полученные результаты сравнили с исправлениями ассессоров, сделанными с учетом контекста запроса (гл. 2.1). Если исправления совпадали, мы считали пару ошибка-исправление контекстно-независимой (например, аналитик: *рюкзк* → **рюкзак**; ассессор: *рюкзк* → **рюкзак**). Если исправления не совпадали, мы считали пару ошибка-исправление контекстно-зависимой (например, аналитик: *крепк* → **скрепка**; ассессор: *крепк* → **крепко**).

Контекстно-независимых исправлений оказалось 74%, что подтвердило наше предположение об однозначности исправлений для большинства словарных ошибок. Примеры исправлений тривиальных ошибок: *актрисса* → **актриса**, *спаисбо* → **спасибо**, *букенестический* → **букинистический**.

¹ Орфоварианты — слова, имеющие одно и то же значение, но различное, как правило, очень близкое, написание (напр. *кэйтлин* → **кейтлин**).

² Неправильная словоформа — слово с опечаткой и его исправление являются разными формами одного и того же слова (напр. *лето* → **летом**).

Среди контекстно-зависимых исправлений большую часть составляют ошибки в коротких словах, допускающие разные, зависящие от контекста, исправления. Например *сво законов* → *свод законов*, и *сво игра* → *своя игра*. Примеры других контекстно-зависимых ошибок представлены в Таблице 4.

Таблица 4. Примеры контекстно-зависимых опечаток

Опечатка	Исправление без контекста	Исправление в контексте запроса
Крепк	Скрепка	<i>крепко</i> заваренный чай
Скчать	Скачать	как не <i>скачать</i> в отпуске
Моне	Моне	эдуард <i>мане</i>

Таким образом, простой эксперимент показал, что для большей части словарных ошибок (74%) исправления в контексте и без контекста совпадают. Это позволяет сделать оценку доли словарных ошибок, которые можно исправлять без учета контекста.

3. Описание подхода

Выбрав класс словарных ошибок в качестве целевого и убедившись, что большая часть словарных ошибок не зависят от контекста, мы можем перейти собственно к оценке надежности исправлений ошибок.

В данной работе мы решаем задачу оценки надежности исправления словарных ошибок как задачу бинарной классификации. Все пары «опечатка-исправление» ($q \rightarrow c$) должны быть отнесены к одному из классов: «надежная» или «ненадежная». Пара ($q \rightarrow c$) считается надежной, если c является правильным исправлением q .

Задача оценки надежности состоит из 4-х этапов:

1. **Исправление.** Для поиска и исправления опечаток в запросах применяется поисковый спелл-чекер Яндексa, определяющий лучший вариант исправления C' запроса Q с помощью модели канала с ошибками [3, 5, 6]:

$$C' = \operatorname{argmax} P(Q|C) \cdot P(C),$$

где $P(Q|C)$ — вероятность трансформации запроса C в запрос Q (модель ошибок), $P(C)$ — вероятность запроса C (языковая модель).

2. **Выравнивание.** Для выделения слов с ошибками пара запрос-исправление ($Q \rightarrow C$) выравнивается по словам (см. гл. 2.1). В результате получаем набор пар опечатка-исправление $\{q_i \rightarrow c_i\}, i = 1 \dots m$.
3. **Фильтрация.** Для выделения ошибок типа «ошибки в отдельных словах» применяются простые фильтры:

- Слово с опечаткой и его исправление не содержат пробелов.
- Слово с опечаткой и его исправление принадлежат одному алфавиту.
Дополнительно отфильтровываются ошибки типа «неправильная словоформа» и ошибки в коротких словах (<4 букв).

4. Классификация. Для каждой пары опечатка-исправление ($q \rightarrow c$) решается задача классификации “надежное/ненадежное” исправление. В качестве метода машинного обучения используется логистическая регрессия.

4. Метрики качества

Эффективность метода определения надежности исправлений оценивается по полноте и точности классификатора:

- Полнота — отношение количества верно классифицированных надежных исправлений ко всем надежным исправлениям.
- Точность — отношение количества верно классифицированных надежных исправлений ко всем исправлениям, которые классификатор отнес к надежным.

5. Признаки

В этой работе при построении классификатора надежности опечаточных исправлений мы использовали признаки только словарного уровня. Были использованы статистические и лексические признаки, широко применяемые в задачах компьютерной лингвистики [14, 12].

1. Вес по словарной языковой модели для q и c . Для вычисления признака использовалась 3-грамная языковая модель, собранная из запросов к поисковой системе Яндекс за полгода [4].
2. Вес по буквенной языковой модели для q и c . Признак вычислялся по 3-грамной буквенной языковой модели, построенной по тем же данным, что и словарная модель.
3. Длины слов q и c в символах.
4. Присутствие q и c в словарных источниках. Слова проверялись по русскому и английскому морфологическим словарям, используемым поисковой системой Яндекс.
5. Язык запроса. Так как в рассматриваемых запросах преобладают русский и английский языки (см. табл. 2), в данной работе использовался бинарный признак en/ru. Язык слова определялся по принадлежности к кириллическому/латинскому алфавиту. Напомним, что q и c принадлежат одному алфавиту (см. гл. 3)

6. Вероятность написания q и c с заглавной буквы. Как показывает опыт, значительную часть плохих подсказок составляют имена собственные, например, редкие фамилии, названия небольших фирм и т. п. Поскольку имена собственные чаще пишутся с заглавной буквы, вероятность написания слова с заглавной буквы можно использовать в качестве признака. Для вычисления этого признака мы использовали корпус, состоящий из 100 миллионов веб-документов.
7. Взвешенная дистанция редактирования для пары $(q \rightarrow c)$. Под дистанцией редактирования понимается вероятность трансформации $q \rightarrow c$, вычисляемая с использованием модели ошибок, описанной в [8].
8. Взаимный словарный контекст q и c . Для q и c строятся вектора слов, с которыми q и c встречаются в 3-грамной словарной языковой модели на расстоянии одного слова. Полученные вектора используются для оценки меры схожести словарных контекстов q и c . В данной работе мы использовали меру схожести контекстов двух слов w_1 и w_2 , предложенную в [10]:

$$P_c(w_2|w_1) = \sum (P(w|w_1) \cdot P(w|w_2) \cdot P(w_2)) / P(w)$$

где w принадлежит множеству слов, которые встречаются как со словом w_1 , так и со словом w_2 . Признак вычислялся по словарной языковой модели.

6. Эксперименты

6.1. Тестовый и обучающий наборы

Для настройки и тестирования предложенного метода были подготовлены обучающий и тестовый наборы. Из дневного лога поисковой системы Яндекс случайным образом были выбраны 30 000 запросов и проверены поисковым спелл-чекером. Спелл-чекер обнаружил и исправил опечатки в 2134 запросах. Для обучения словарного классификатора мы извлекли из пар запрос-исправление все пары слово-исправление. Таких пар получилось 2545. Из них мы удалили пары, состоящие из коротких слов (354 пары) и орфовариантов (41 пара), после чего оставшиеся 2150 пар были переданы на разметку аналитику. Разметка состояла в том, чтобы разделить пары слово-исправление на правильные и неправильные. Правильных оказалось 77%. Размеченный набор пар слово-исправление был разделен на две выборки (обучающую и тестовую), равные по размеру и по содержанию правильных и неправильных исправлений.

6.2. Базовый набор признаков

За базовый уровень мы приняли показатели классификатора, построенного на признаках, используемых спелл-чекером для выявления и исправления ошибок в запросах. Это веса слова и его исправления по языковой модели, а также взвешенная дистанция редактирования слова и исправления (см. признаки 1,7 в гл. 5).

6.3. Результаты

Учитывая тот факт, что разные поисковые задачи предъявляют разные требования к точности классификации надежности исправлений, для оценки возможности настройки классификатора мы проводили эксперименты с двумя порогами точности (0,9 и 0,95). Результаты, полученные на тестовой выборке, представлены в Таблице 5:

Таблица 5. Качество классификатора на тестовой выборке

Набор признаков	Полнота (порог точности 0,9)	Полнота (порог точности 0,95)
Базовый	0,364	0,214
Полный	0,773	0,549

По результатам эксперимента можно сделать следующие выводы:

Полный набор обеспечил прирост полноты автозамен на $0,3 \div 0,4$, что свидетельствует о высокой эффективности признаков словарного уровня.

Если учесть, что опечатки составляют 63,7% всех ошибок (табл. 3), полнота классификатора 0,549 означает возможность автоматически исправлять 35% словарных ошибок в запросах с точностью не ниже 0,95. Это весьма высокий показатель для функции автоматической коррекции. Предложенный метод можно признать пригодным и рекомендовать для автоисправления запросов в поисковой системе.

Из анализа результатов эксперимента следует, что наиболее эффективным признаком оказался «взаимный словарный контекст». Особенность этого признака заключается в том, что он работает не с контекстом данного запроса, а агрегирует информацию о взаимозаменяемости слов в одних и тех же контекстах по разным запросам, собранным за большой период времени. Эта информация позволяет с высокой точностью установить, является ли одно слово опечаткой другого. К сожалению, этот признак не работает для редких слов (из-за отсутствия статистических данных об их взаимных контекстах).

Среди ошибок первого рода следует указать неверное отнесение редких слов к классу опечаток. Проблема заключается в нехватке статистических данных для построения значимых признаков (взаимный контекст, вероятность написания с заглавной буквы) для редких слов (фамилии, названия небольших компаний, новых сайтов и т. п.). Здесь необходимо увеличивать объемы исходных данных, а также добавлять новые признаки, например, присутствие слова в специализированных словарях (товаров, топонимов и т. п.).

Ошибки второго рода связаны с признаком присутствия слов в морфологическом словаре. Для исправления контекстно-зависимых ошибок у классификатора словарного уровня недостаточно данных. В этих случаях, очевидно, необходимы контекстные признаки, не используемые в данной работе.

7. Заключение

В данной работе решалась проблема повышения эффективности автоматического исправления ошибок в поисковых запросах. В качестве целевого класса были выбраны словарные ошибки (пропуск/вставка/замена/перестановка букв в словах), составляющие две трети всех ошибок в запросах. Было показано, что значительная часть ошибок являются тривиальными (исправление очевидно и однозначно), не зависят от словарного окружения и могут быть исправлены в автоматическом режиме.

Для принятия решения о возможности автоматического исправления использовался бинарный классификатор, разделяющий исправления на надежные, пригодные для автозамен, и ненадежные, пригодные только для подсказок. Поскольку тривиальные ошибки не зависят от контекста, для определения надежности исправлений достаточно признаков словарного (бесконтекстного) уровня, что значительно упростило задачу подбора признаков. В работе были использованы наиболее распространенные лексические и статистические признаки, применяемые при решении поисковых и лингвистических задач. Построенный на их базе классификатор показал приемлемое качество и возможность регулирования баланса полнота/точность. С помощью предложенного в работе метода можно с высокой точностью автоматически исправлять больше половины словарных опечаток, т. е. почти треть всех ошибок в поисковых запросах.

Несмотря на положительные результаты, работу по повышению качества автоматического исправления ошибок можно продолжить по многим направлениям. Помимо добавления новых признаков и исправления «несловарных» типов ошибок, в качестве перспективных задач можно назвать добавление новых классов надежности подсказки и использование данных о пользовательских кликах.

References

1. *Baba Y., Suzuki H.* (2012) How Are Spelling Errors Generated and Corrected? A Study of Corrected and Uncorrected Spelling Errors Using Keystroke Logs. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, pages 373–377, Jeju, Republic of Korea, 8–14 July 2012.
2. *Baldwin T., Chai J. Y.* Autonomous Self-Assessment of Autocorrections: Exploring Text Message Dialogues. In: 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 710–719, Montreal, Canada, 2012.
3. *Bajtin A.* (2008) [Ispravlenie poiskovyh zaprosov v Yandekse. Veroyatnostnaja jazykovaja model] Rossijskie internet-tehnologii 2008.
4. *Brants T., Popat A. C., Xu P., Och F. J., Dean J.* (2007). Large Language Models in Machine Translation. In: EMNLP'07. pp. 858–867
5. *Brill E., Moore R. C.* (2000). An Improved Error Model for Noise Channel Spelling Correction. In: ACL'00. pp. 286–293
6. *Cucerzan S., Brill E.* (2004). Spelling Correction as an Iterative Process that Exploits the Collective Knowledge of Web Users . In: EMNLP'04 pages 293–300
7. *Erehinskja T. N., Titova A. S., Okat'ev V. V.* (2011) Syntax Parsing For Texts With Misspellings In Dictascope Syntax [Sintaksicheskij analiz teksta s orfograficheskimi oshibkami v sisteme Dictascope Syntax] Trudy Mezhdunarodnoj Konferentsii “Dialog 2011” [Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog 2011”], Bekasovo.
8. *Karpenko M. P., Protasov S. V.* (2011) Some Methods for Language Model Pruning [Nekotorye metody ochistki slovarja zaprosov poiska] Trudy Mezhdunarodnoj Konferentsii “Dialog 2011” [Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog 2011”], Bekasovo.
9. *Kukich K.* (1992) Techniques for Automatically Correcting Words in Text. In: ACM Computing Surveys, Vol. 24, No. 4.
10. *Li M., Zhu M., Zhang Y., Zhou M.* (2006). Exploring Distributional Similarity Based Models for Query Spelling Correction . In: ACL'06. pp. 1025–1032.
11. *Rabiner, L. R.* (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. In: Proceedings of the IEEE, 77 (2), pp. 257–286
12. *Specia L., Cancedda N., Dymetman M., Turchi M., Cristianini, N.* (2009). Estimating the Sentence-Level Quality of Machine Translation Systems. In: EAMT'09. pp. 28–35
13. *Specia L., Hajlaoui N., Hallett C., Aziz W.* (2011). Predicting Machine Translation Adequacy. In: MTSummit'11. pp. 513–520
14. *Whitelaw C., Hutchinson B., Chung G. Y., Ellis G.* (2009). Using the Web for Language Independent Spellchecking and Autocorrection. In: EMNLP'09. pp. 890–899
15. http://productforums.google.com/forum/#!topic/websearch/g_XQupJ8Pvgcc