

## Формирование поискового запроса для поиска информации в предметной области с применением Закона Ципфа и правила трёх сигм

*И.С. Вакурин<sup>1</sup>, Л.А. Трemasова<sup>1</sup>, А.В. Алёшинцев<sup>2</sup>, Д.В. Гадасин<sup>1</sup>*

<sup>1</sup>Московский технический университет связи и информатики, Москва

<sup>2</sup>Московский государственный строительный университет, Москва

**Аннотация:** Ежегодный рост нагрузки на центры обработки данных увеличивается в многократных размерах, что обусловлено возрастающим ростом пользователей информационно-телекоммуникационной сети Интернет. Пользователи обращаются к различным ресурсам и источникам, используя для этого поисковые системы и сервисы.

Установка оборудования, которое обрабатывает телекоммуникационный трафик быстрее, требует значительных финансовых затрат, а также может существенно увеличить время простоя центра обработки данных за счет возможных проблем во время проведения регламентных технических работ. Целесообразнее сосредоточить ресурс на улучшении программной, а не аппаратной части оборудования.

В статье приводится алгоритм, за счёт которого возможно снизить нагрузку на телекоммуникационное оборудование за счет поиска информации внутри определённой предметной области, а также путём использования особенностей естественного языка и процесса образования в нём слов, предложений и текстов.

Анализировать запрос предложено на основании формирования префиксного дерева и кластеризации, а также путём расчета вероятности возникновения нужного слова на основе правила трёх сигм и Закона Ципфа.

**Ключевые слова:** правило трёх сигм, Закон Ципфа, кластеры, анализ языка, морфемы, префиксное дерево, распределение вероятностей.

**Введение.** В предметной области, которая описывают процессы хранения обработки и передачи информации, возможно определить понятие информационного поля. Информационное поле подразумевает, прежде всего, набор элементов, между которыми существует система. Один из видов взаимодействия в рамках информационного поля между субъектами для передачи информации предполагает использование языковых форм, которые являются неотъемлемой составляющей речи.

В поле для взаимосвязи и переходах между элементами определены операции сложения, инверсии, умножения и деления. В русском языке для сложения используется союз "и", он часто используется для перечисления, тем самым составляя некоторую алгебраическую прогрессию. Роль противоположного выполняют антонимы. Умножение и деление служат для

определения множеств элементов в поле, математически это можно описать при помощи рядов, можно свернуть и развернуть суммы в нём [1-3].

Информационное поле можно разделить на группы по языковому признаку, где детерминантой группы является язык взаимодействия. Язык состоит из лингвистических единиц. Под лингвистическими единицами понимаются морфемы, слога, слова, словосочетания, предложения и т.д..

Совокупность единиц, возможно, отнести к нечетким множествам, т.к. заранее точно не известно какое подмножество сформирует информацию необходимую для передачи. Составные части языка имеют достаточно слабую зависимость друг от друга, но при этом ни один фактор или признак у них не является доминирующим, что позволяет сказать, что они подчиняются условиям Центральной предельной теоремы (ЦПТ). Согласно ЦПТ, такие множества, возможно описать при помощи закона нормального распределения, так как, большое множество слабосвязанных величин одинакового масштаба имеют распределение близкое к нормальному.

Информация, сформированная из одних и тех же элементов может находиться близко друг к другу (сильносвязная), если она относится к одной и той же предметной области, и на расстоянии друг от друга (слабосвязная), если относятся к разным предметным областям, которые к тому же и противоположны относительно смысловой нагрузки. Например, слово "ключ" может применяться в предметной области замков, дверей или криптографии, как предмет для доступа к чему-либо, а может применяться, как описание водного источника в природе. Если предметные области замков и дверей находятся на близком расстоянии друг от друга, криптография и водные ресурсы достаточно далеки друг от друга. В такой ситуации, возможно, использовать закон больших чисел, который применим к зависимым случайным величинам, если только сильная зависимость существует между случайными величинами с близкими номерами, а между

---

случайными величинами с далекими номерами зависимость достаточно слаба. Если между элементами, например, группы слов в определенной предметной области существует большая взаимосвязь, то эти слова формируют математическое ожидание и будут встречаться в тексте на определённую тематику чаще.

Исходя из предположения, что морфемы русского языка имеют нормальное распределение, то для них возможно использовать правило трех сигм, которое определяет, что для любой случайной величины с конечной дисперсией вероятность того, что случайная величина отклонится от своего математического ожидания не менее, чем на три среднеквадратических отклонения, т.е. лежит в интервале  $[\mu - 3\sigma; \mu + 3\sigma]$  изображенные на рис.1.

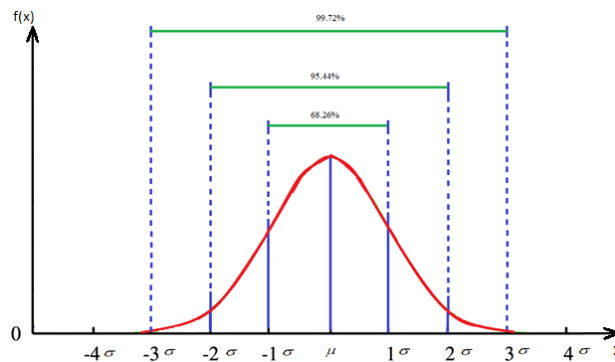


Рис. 1. – Правило трех сигм ( $\mu$  - математическое ожидание,  $\sigma$  - дисперсия)

Из рисунка видно, что в интервале  $[\mu - \sigma; \mu + \sigma]$  находится 68.26% значений, которые принимает случайная величина. Процент значений, которые принимает случайная величина, увеличивается при изменении интервала и в интервале  $[\mu - 3\sigma; \mu + 3\sigma]$  находится 99.72% из всех морфем, т.е. случайная величина примет значение, с отклонением более трех сигм только для 0.28%, что само по себе крайне мало. Следовательно, в зависимости от вероятности появления морфемы в тексте, возможно определить, в какой именно интервал попадет морфема.

**Применения правила трёх сигм для слов русского языка.** Исходя из того, что понятие «слово» является важнейшей структурно-семантической

единицей языка, которая служит для наименования предметов, процессов, свойств, отличается самостоятельностью и свободным воспроизведением в речи, и представляет собой основу для предложения, то процесс словообразования в русском языке, возможно, представить исходя из рис. 2.

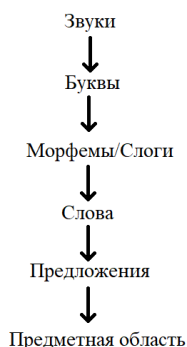


Рис. 2. – Словообразование в русском языке

Можно выделить шесть основных уровней, на самом нижнем уровне находятся абстракции в соответствии которым субъект (исходя из своих собственных знаний) подбирает звуки, звуки также являются абстракциями и для каждого звука, возможно, поставить в соответствие буквы. Путём группирования букв получают слога или морфемы, а их совокупность уже образует слово. Множество слов возможно преобразовать в предложения или просто оставить в виде упорядоченного множества, что служит для описания окружающей действительности (предметная область). Таким образом, наиболее устойчивыми частями языка являются морфемы / слога, которые видоизменяют слова, прирастая к ним и изменяя их форму.

На третьем уровне находятся морфемы и слоги. Они имеют сходства, но гораздо важнее их отличия. Слоги строятся из соображений фонетики и состоят из звуков, чья природа имеет наибольшую акустическую гармонию, например, сочетание звуков «ся», «ля», «ом» звучат гармонично и легко произносятся, когда сочетания звуков «нм», «рн» и «вц» произнести сложно, поэтому они и не являются слогами. Слог никак не отражает смысловое

формирование слова, формируясь только с позиции эстетики и эргономики, потому и играет обширную роль в поэзии и музыке.

Формирование слов из морфем, в свою очередь, имеет как раз большую смысловую основу. Морфема представляет собой некую абстракцию, считаясь не знаком, а скорее классом знаков. Реализация морфемы при письме называется морфом, а совокупность морфов одной морфемы с одинаковым звучанием алломорфом [4]. Например, в предложении: «Я бегу, и ты бежишь, а он не бежит» морфема «бег-» есть в трёх формах: 1. «бег-» в «бегу»; 2. «беж-» в «бежишь»; 3. «беж» в «бежит». Но алломорфов в нём два: «бег-» и «беж-». Чтобы два морфа относились к одному алломорфу они должны обязательно иметь одинаковый фонемный состав и ударение.

Например, существует слово «ход». На рис. 3 отображён процесс его изменения.

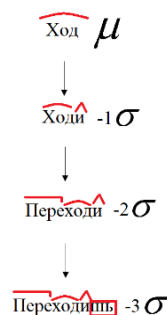


Рис. 3. – Процесс изменения слова ( $\mu$  - мат. ожидание,  $\sigma$  - дисперсия)

Корень слова является в данном случае математическим ожиданием. Все приращения, к слову, являются дисперсией, и порядок дисперсии увеличивается при каждом новом приращении.

На первой сигме располагается суффикс, так как суффикс является самой распространённой морфемой в русском языке и статистически встречается чаще всего и обычно служит для перехода слова от существительного к другим частям речи. Всего суффиксов в русском языке около шестисот.

На втором удалении приставка, их в русском языке около 60. Чаще всего они обозначают преодоление чего-то или принадлежность существительного к чему-либо.

На третьем месте по частоте морфем – окончания. Их немного меньше приставок, около 50. Чаще всего они указывают на что-то и служат для связи слов и словосочетаний. Такая схема может быть представлена при помощи графика плотности вероятности. На рис. 4 показан график плотности вероятности для слова «Ход».

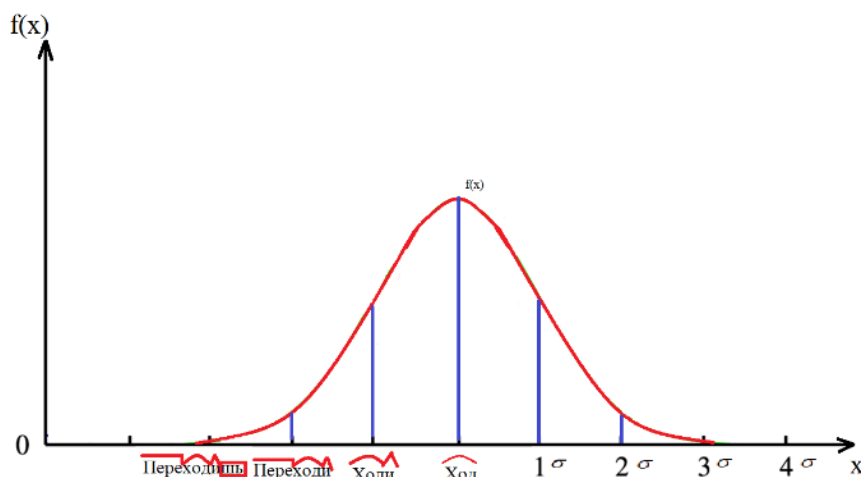


Рис. 4. – График плотности вероятности для слова «Ход» ( $\sigma$  - дисперсия)

Как можно понять из графика, чем больше морфем содержит слово, тем реже оно встречается в письме или речи. Например, если мы обсуждаем предметную область часовых механизмов, то возникновения в ней слова «ход» достаточно велико, а в зависимости от контекста слово может при помощи морфем видоизменяться. Формула (1) показывает стандартное выражение для математического ожидания в нормальном распределении:

$$M(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx, \quad (1)$$

где:  $x$  - случайная величина;  $f(x)$  - функция распределения случайной величины.

Функция распределения из формулы (2):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (2)$$

где:  $\sigma$  - дисперсия;  $\mu$  - математическое ожидание;  $x$  - случайная величина.

Функция распределения из (2) даёт представление о том, как часто в языке встречается конкретная морфема. При помощи языка можно описывать окружающие явления, при этом при описании каждого явления будут использоваться одни и те же морфемы, но вероятность их появления будет отличаться [5]. Можно построить график для любой предметной области, который отразит вероятность появления морфем, а предметную область представить в виде кластера морфем. На рис. 5 показано, как распределяются морфемы в областях отклонения в определенной предметной области.

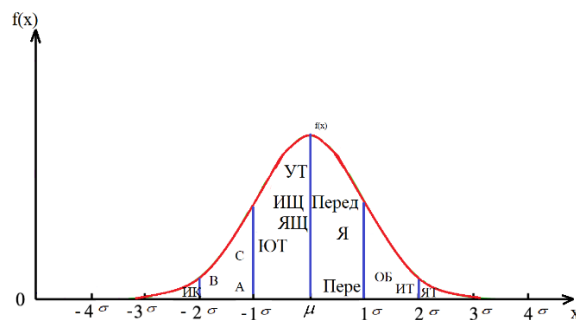


Рис. 5. – Распределение морфем по областям дисперсии ( $\sigma$  - дисперсия)

Морфемы, находящиеся в первом интервале, появляются в текстах предметной области наиболее часто, соответственно морфемы из второго интервала появляются реже, а из третьего интервала совсем редко. Это может быть обусловлено стилистическим окрасом текста, так как обороты в художественном и научно-публицистическом стиле достаточно сильно различаются между собой [6]. Например, уменьшительно-ласкательные суффиксы в художественном тексте будут встречаться гораздо чаще.

Для любого кластера характерен центроид, который взаимодействует с другими элементами кластера, при этом возможно взаимодействие элементов кластера между собой. Такое взаимодействие или взаимосвязь графически

выражается в виде направленного графа, где направления формируют пути формирования слов из морфем (из слов предложений) посредством которых описывается данная предметная область [7]. Центроид кластера должен входить в интервал  $[\mu - \Delta; \mu + \Delta]$ , где  $\mu \pm \Delta$ ,  $\mu \pm \sigma$  соответственно. В таком случае слово «ход» будет представлять собой центроид, а в графовом представлении в виде дерева будет иметь индекс 0, что соответствует значению математического ожидания для графика на рис. 6.

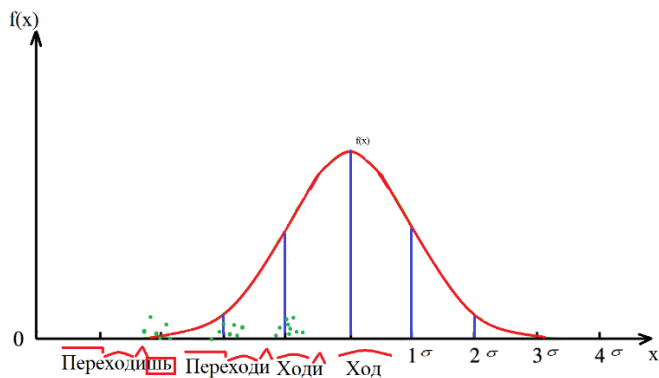


Рис. 6. – Кластеры на графике плотности вероятности

**Морфемные кластеры.** Слова в русском языке могут изменяться по родам, числам, падежам и приобретать различный эмоциональный окрас, в таких изменениях непосредственное участие принимают морфемы. Сами по себе такие преобразования (без привязки к конкретному контексту) образуют ряд. Например, как отдельные числа на числовой прямой, с помощью морфем происходит взаимоувязка отдельных слов в единое целое, образующее осмысленное описание предметной области. Производится группировка морфем в окрестностях каждой сигмы (области отклонения) и для каждой морфемы определяется вероятность ее появления. Соответственно, в каждой области отклонения возможно сформировать группы морфем или морфемные кластеры. Морфемные кластеры представляют собой абстрактный набор всех приставок, суффиксов и окончаний, существующих у слова. Морфемы группируются в кластеры по общему признаку в каждой области отклонения: суффиксы, приставки и окончания, показанные на рис 6.



Более детально на примере приставок это отображено на рис. 7.

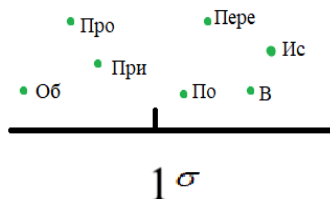


Рис. 7. – Кластер приставок ( $\sigma$  - дисперсия)

За центроид кластера принимается значение сигмы. Вид и размер кластеров зависят от исходного слова и меняются в зависимости от результата на каждом шаге, так как набор группы на следующем шаге зависит от выбранной морфемы на текущем шаге, исходя из того, что сочетания одних морфем в языке встречаются чаще других.

В исходном виде кластер представляет собой некоторую структуру с общей точкой притяжения. Более упорядоченно кластер можно представить в виде префиксного дерева, листья которого состоят из одномерного массива и имеют два элемента: значение и ключ. Путь от корня дерева до листьев его нижнего уровня определяет формирование морфемы, в том случае если формирование морфемы не возможно, то ключ в листе отсутствует, т.е. в листьях дерева содержатся значения индекса, из которых формируются ключи, а корни значений не содержат. Деревья формируются для каждой из сигм. На рис. 8 изображено префиксное дерево для приставок на букву «п».

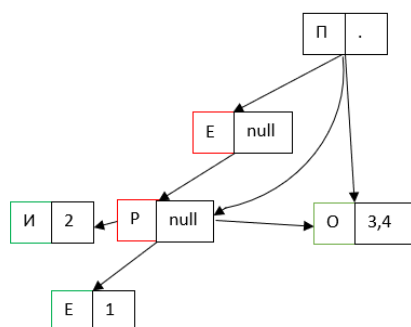


Рис. 8. – Префиксное дерево с корнем «п»

Исходя из дерева рис. 8 можно получить значения для ассоциативного массива, представленные в таблице № 1. В листьях со значением null нет

никаких ключей, так как в русском языке нет приставок «пе» или «пер». Листу «о» соответствует два значения индекса, что означает, что лист представлен в двух путях, в нашем случае это путь «ПО» и «ПРО».

Таблица № 1

Значения ассоциативного массива

Индекс	Ключ	Индекс	Ключ
Пере	1	Про	3
При	2	По	4

Для формирования возможного слова необходимо выбрать корень случайным образом, и при последовательном его прохождении через дерево приставок, суффиксов и окончаний (порядок строго закреплен) определяется ассоциативный массив [8]. Формирование слова при прохождении через морфемные кластеры показано на рис. 9.



Рис. 9. – Формирование слова при прохождении через морфемные кластеры

Исходя из того, что морфемы уникальны и получаются из соответствующих деревьев, морфему определить посредством индекса. Последовательность индексов морфем определяет ключ, который возможно использовать для поискового запроса.

Следующий этап заключается в формировании запроса, состоящего из последовательности слов. Для алфавита языка, для каждого его символа, возможно, определить вероятность её возникновения в тексте, например в

таблице № 2 представлены вероятности появления букв в тексте русского языка.

Таблица № 2

Вероятность (В) возникновения букв в тексте (Символ (С))

С	В	С	В	С	В	С	В
Пробел	0,175	Р	0,040	Я	0,018	Х	0,009
О	0,090	В	0,038	Ы	0,016	Ж	0,007
Е	0,072	Л	0,035	З	0,016	Ю	0,006
А	0,062	К	0,028	Ъ	0,014	Ш	0,006
И	0,062	М	0,026	Б	0,014	Ц	0,004
Н	0,053	Д	0,025	Г	0,013	Щ	0,003
Т	0,053	П	0,023	Ч	0,012	Э	0,003
С	0,045	У	0,021	Й	0,010	Ф	0,002

Следовательно, в одной предметной области, для которой определено текстовое описание, существует некоторая вероятностная прогрессия по аналогии с рис. 9, начинающаяся с буквы и заканчивающаяся текстом. Все этапы показаны на рис. 10.

Буква → Слово → Предложение → Текст

Рис. 10. Этапы вероятностной прогрессии

Рассмотрим процесс формирования предложения «Переходишь улицу в неполюженном месте» в предметной области правил дорожного движения. Каждый текст в любой предметной области начинается с буквы. После буква преобразуется в слово с использованием префиксного дерева. Данный процесс показан на рис. 11.



Рис. 11. – Этап формирования слова

Первая буква, введенная в предложении, становится префиксом, за ней следует вторая буква, то есть индекс, но в следующей итерации она снова

уже становится префиксом. В тексте между словами разделителем является символ пробел. Именно по пробелу определяется, где останавливается формирование префиксного дерева и таблицы соответствия [9].

В следующем этапе слово начинает преобразовываться для встраивания в предложение, то есть обрастает морфемами, данный процесс показан на рис. 9. После того, как слово обросло морфемами, она встаёт на своё место в предложение. Предложение так же формируется при помощи префиксного дерева. Процесс формирования предложения показан на рисунке 12.



Рис. 12. – Процесс формирования предложения

Всего в таком предложении получается три перехода, без учета начального перехода от первого слова.

**Условная вероятность в префиксном дереве.** Исходя из того, что предметная область выражается посредством префиксного дерева, в нём, возможно, организовать поиск информации. Поиск представляет собой один из способов упорядоченного обхода дерева – алгоритм поиска. Алгоритм может быть, как детерминированным, результат поиска точно совпадает с условиями поиска, так и вероятностным, результат поиска не в полной мере совпадает с поисковым запросом. В случае вероятностного поиска, возможно говорить о погрешности поиска, которая определяет его точность. Все пространство, в котором проводится поиск, делится между тремя сигмами на области: тематическая, смежная тематика, область наименьших совпадений. Изначально слова и морфемы распределяются на три области (области

отклонений) [10]. После распределения между областями устанавливается процентное соотношение, показано на рис. 13 согласно правилу трех сигм:  $\sigma_1$  – область, тематическая – 68,2%;  $\sigma_2$  – область, смежные тематики – 27,2%;  $\sigma_3$  – область, маловероятные пересечения – 4,2 %.

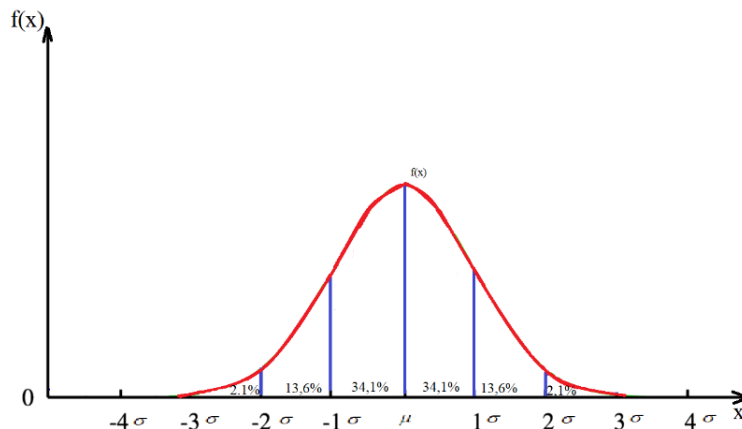


Рис. 13. – Правило трех сигм с разбиением по областям  $\sigma$  (дисперсия)

В первой области находятся слова или части слов, которые являются характерными для предметной области, например, в предметной области ПДД может часто встречаться слово улица, что увеличивает его частоту появления в тексте. Во второй области могут быть слова из тем, которые так или иначе с темой ПДД имеют пересечения. Например, слово улица может употребляться не только в контексте ПДД, но и в контексте городского планирования, что, в конечном счете, влияет на ПДД. В границах третьей области расположены слова, которые имеют скорее случайное пересечение с контекстом документа, например, слово улица встречается в стихотворных произведениях Александра Блока, но улица там упомянута в художественном, а не прикладном смысле. Оставшиеся 0,04% попадают в область маловероятной статистической погрешности за пределами трёх областей отклонения.

Несмотря на то, что все три области могут содержать одинаковые слова, но поскольку алгоритм ориентируется на конкретную тематику, то вероятность подстановки слова «проезжая» после слова «улица» будет

больше, чем подстановка слова «фонарь». В данном случае уместно говорить о таком понятии, как условная вероятность, т.е. при условии появления слова А в контексте предметной области, с некоторой вероятностью за ним или перед ним появится конкретное слово В. Такая вероятность рассчитывается по формуле Байеса приведена формула (3):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad (3)$$

где:  $P(A)$  – априорная вероятность;  $P(B|A)$  – вероятность наступления события В при наступившем событии А;  $P(A|B)$  – апостериорная вероятность наступления события А при наступлении события В;  $P(B)$  – полная вероятность наступления события В.

Событие А соответствует действию ввода любого слова. Событие В соответствует действию по формированию поискового запроса, т.е. вводу последнего слова в поисковом запросе, допускается что запрос может состоять из одного слова, но вероятность данного события мала.

При формировании поискового запроса алгоритм рассчитывает вероятность появления следующего слова, то есть наступление события В, при наступлении события А. Например, при начальном формировании поискового запроса и ввода слова «переходишь», событие А наступит с вероятностью, близкой к единице, так как запрос только формируется и в областях отклонения, находятся слова, которые могут появляться после ввода слова «переходишь». Например, с вероятностью 68% после слова «переходишь», будет находиться слово «улицу», с вероятностью 26% будет находиться слово «к», а с вероятностью 4% будет слово «через». Вероятность появления следующего слова зависит не только от той области, в которой оно находится, но и от ранга слова внутри ее. Ранг слова в предметной области (согласно Закону Ципфы) зависит от частоты его упоминания. Закон определяется формулой (4). и согласно ему второе по используемости слово

---

будет встречаться примерно в 2 раза реже, чем первое, третье – в три раза реже, чем первое, и далее:

$$N_n = N_a / n, \tag{4}$$

где:  $N_n$  – количество слов определенного ранга;  $N_a$  – число всех слов;  $n$  – ранг слова.

Ранг слова представляет собой порядковый номер слова, определяем его частотой, описывается распределением Парето. Таким образом вероятность события В зависит от ранга слов, входящих в поисковый запрос, а ранг в свою очередь от частоты употребления слов в предметной области. На рис. 14 показана схема описанного алгоритма.

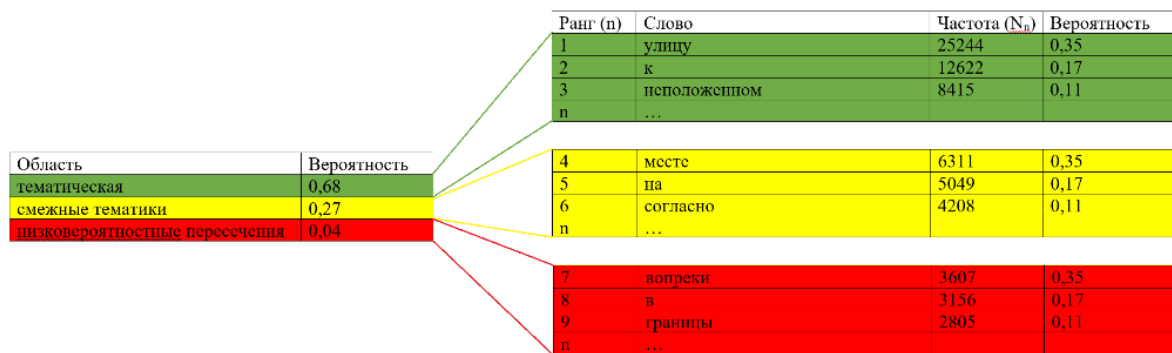


Рис. 14. – Схема алгоритма

При формировании поискового запроса и ввода первого слова алгоритм вначале обращается в область отклонения, где с вероятностью в 68,8% находится нужное слово, и только потом в области смежных тем и низковоероятностных пересечений. Далее внутри области производится обращение к рангу слова для составления поисковой выдачи. Ранжирование в каждой области отклонения начинается заново. Пример поисковой выдачи в результате работы алгоритма показан на рис. 15.

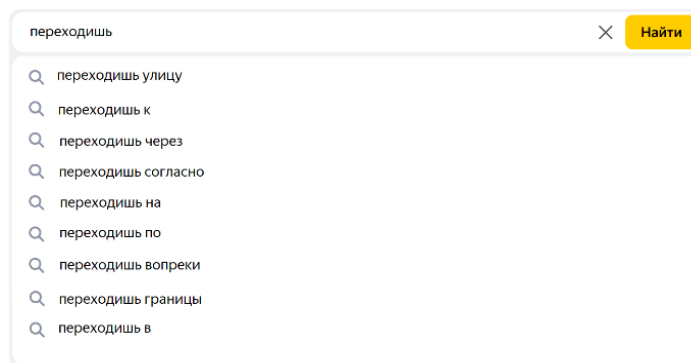


Рис. 15. – Пример поисковой выдачи по алгоритму

Для расчета вероятностей в алгоритме используются формула расчета полной вероятности формула (5) и формула (6) расчета зависимости события В от события А:

$$P(B) = P(B | A)P(A) \quad (5)$$

$$P(B | A) = \frac{P(A | B)P(B)}{P(A)} \quad (6)$$

Для расчета полной вероятности формирования запроса понадобится несколько итераций, количество которых будет зависеть от количества слов в конечном запросе. В запросе «Переходишь улицу в неполюженном месте» помимо первого слова вводимого с клавиатуры, есть ещё четыре слова, значит нужно будет 4 повторения, как было сказано выше вероятность события А стремится к 1, возьмем  $P(A)$  равным 0.98,  $P(A|B)$  равным 0.35 а  $P(B)_1$  равным 0.68, так как слово «улицу» находится в первой области отклонения и идёт первым по рангу, рассчитаем  $P(B|A)$  и  $P(B)$  для первой итерации по формулам (7) и (8), соответственно:

$$P(B | A) = \frac{0,35 \cdot 0,68}{0,98} = 0,19 \quad (7)$$

$$P(B)_1 = 0,19 \cdot 0,98 = 0,19 \quad (8)$$

Слово «в» находится в третьей области отклонения, его параметры рассчитываются по формулам (9) и (10):



$$P(B|A) = \frac{0,04 \cdot 0,17}{0,19} = 0,03 \quad (9)$$

$$P(B)_2 = 0,03 \cdot 0,18 = 0,07 \quad (10)$$

Слово «неположенном» находится в первой области отклонения, его параметры рассчитываются по формулам (11) и (12):

$$P(B|A) = \frac{0,11 \cdot 0,68}{0,07} = 0,98 \quad (11)$$

$$P(B)_3 = 0,98 \cdot 0,068 = 0,07 \quad (12)$$

Слово «месте» находится во второй области отклонения, его параметры рассчитываются по формулам (13) и (14):

$$P(B|A) = \frac{0,27 \cdot 0,35}{0,07} = 0,42 \quad (13)$$

$$P(B)_4 = 0,42 \cdot 0,07 = 0,03 \quad (14)$$

Полная вероятность формирования поискового запроса рассчитывается по формуле (15):

$$P(B) = P(B)_1 + P(B)_2 + P(B)_3 + P(B)_4 = 0,19 + 0,07 + 0,07 + 0,03 = 0,36 \quad (15)$$

Такой показатель вероятности формирования полного поискового запроса обусловлен наличием слов из 2 и 3 области отклонения в запросе. Для увеличения релевантности выдачи по такому запросу необходимо будет уточнить его.

### Выводы

В результате проведенной работы было описано правило трех сигм применительно к русскому языку на основе Центрально Предельной Теоремы и нормального распределения Гаусса. Описано понятие морфемных кластеров и их взаимосвязь с префиксными деревьями, как основы для построения в языке слов и предложений. На основе правила трех сигм, формулы Байеса и закона Ципфа синтезирован алгоритм построения поискового запроса для поиска информации в предметной области.

---

## Литература

1. Гадасин Д.В., Шведов А.В. Применение транспортной задачи для балансировки нагрузки в условиях нечеткости исходных данных // Т-Comm: Телекоммуникации и транспорт. 2024. Т. 18, № 1. С. 13-20.
  2. Егунова А.И., Комаров Р.С., Федосин С.А., Шибайкин С. Д., Жарков Р.А. Сравнительный анализ методов извлечения знаний из текстов для построения онтологий // Инженерный вестник Дона. 2025. № 2. URL: [ivdon.ru/uploads/article/pdf/IVD\\_60N1y25\\_Egunova\\_Komarov.pdf\\_e44b195e53](http://ivdon.ru/uploads/article/pdf/IVD_60N1y25_Egunova_Komarov.pdf_e44b195e53).
  3. Алёшинцев А.В., Сак А.Н. Моделирование информационных систем компьютерного зрения для машинного перевода методами теории графов // Т-Comm: Телекоммуникации и транспорт. 2018. Т. 12, № 10. С. 58-63.
  4. Гадасин Д.В. Построение бинарного дерева минимальной цены // Т-Comm: Телекоммуникации и транспорт. 2024. Т. 18. № 11. С. 38-44.
  5. Turcotte D.L. Fractals and Chaos in Geology and Geophysics / 2nd ed. Cambridge University Press; 1997. 416 p.
  6. Fischer P. Chaos, Fractals, and Dynamics / Taylor&Francis, 1985. 280 p. URL: [doi.org/10.1201/9781003072676](https://doi.org/10.1201/9781003072676).
  7. Петренко И.А., Гулынина Е.В., Донева О.В. К вопросу практического применения закона больших чисел // Университетская наука. 2017. № 1(3). С. 147-150.
  8. Белов В.В., Чистякова В.И. Алгоритмы и структуры данных: учебник. М.: КУРС: ИНФРА-М, 2023. 240 с.
  9. Аксенов К.А., Сунь Л., Калинин И.А. Разработка системы распознавания намерений и сущностей для систем вопросов и ответов с использованием “no code” платформы "TWIN" // Инженерный вестник Дона. 2024. № 9. URL: [ivdon.ru/uploads/article/pdf/IVD\\_27N9y24\\_Sun\\_Lina.pdf\\_e1d5c9879a.pdf](http://ivdon.ru/uploads/article/pdf/IVD_27N9y24_Sun_Lina.pdf_e1d5c9879a.pdf)
-



10. Stallings W. Cryptography And Network Security. Principles And Practice // 5th Edition, 2011. 719 p.

### References

1. Gadasin D.V., Shvedov A.V. T-Comm : Telekommunikacii i transport, 2024. T. 18, № 1. pp. 13-20. DOI 10.36724/2072-8735-2024-18-1-13-20.
2. Egunova A.I., Komarov R.S., Fedosin S.A., Shibajkin S. D., Zharkov R.A. Inzhenernyj vestnik Dona, 2025. № 2. URL: [ivdon.ru/uploads/article/pdf/IVD\\_60N1y25\\_Egunova\\_Komarov.pdf\\_e44b195e53](http://ivdon.ru/uploads/article/pdf/IVD_60N1y25_Egunova_Komarov.pdf_e44b195e53).
3. Aleshintsev A.V., Sak A.N. T-Comm : Telekommunikacii i transport, 2018. T. 12, № 10. pp. 58-63. DOI 10.24411/2072-8735-2018-10158.
4. Gadasin D.V. T-Comm : Telekommunikacii i transport, 2024. T. 18. № 11. С. 38-44. DOI 10.36724/2072-8735-2024-18-11-38-44.
5. Turcotte D.L. [Fractals and Chaos in Geology and Geophysics]. 2-e izd. Cambridge University Press; 1997. 416 p.
6. Fischer P. Chaos, Fractals, and Dynamics. Taylor&Francis, 1985. 280 p. doi.org/10.1201/9781003072676.
7. Petrenko I.A., Guly`nina E.V., Doneva O.V. Universitetskaya nauka. 2017. № 1(3). pp. 147-150.
8. Belov V.V., Chistyakova V.I. Algoritmy` i struktury` danny`x: uchebnik. Algorithms and data structures: textbook. M.: KURS: INFRA-M, 2023. 240 p.
9. Aksenov K.A., Sun` L., Kalinin I.A. Inzhenernyj vestnik Dona, 2024. № 9. URL: [ivdon.ru/uploads/article/pdf/IVD\\_27N9y24\\_Sun\\_Lina.pdf\\_e1d5c9879a.pdf](http://ivdon.ru/uploads/article/pdf/IVD_27N9y24_Sun_Lina.pdf_e1d5c9879a.pdf)
10. Stallings W. Cryptography And Network Security. Principles And Practice], 5th Edition, 2011. 719 p.

**Дата поступления: 30.12.2024**

**Дата публикации: 3.02.2025**