

УДК 004.584:004.891.2  
DOI: 10.15827/0236-235X.130.276-282

Дата подачи статьи: 10.03.20  
2020. Т. 33. № 2. С. 276–282

## **Контекстное ранжирование ответов по относительным метрикам в диалоговых системах помощи пользователям**

Н.Р. Гафуров<sup>1</sup>, аспирант, [nail@gafurov.ru](mailto:nail@gafurov.ru)

<sup>1</sup> Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики (Университет ИТМО), г. Санкт-Петербург, 197101, Россия

Предметом исследования являются диалоговые системы помощи пользователям веб-приложений или онбординг-системы.

Рассматриваются диалоговые системы в целом, вероятные причины их текущей актуализации, условия целесообразности применения, их виды и обосновывается выбор вида для описываемой системы.

Исследованы системы помощи, порождающие необходимость отвечать на вопросы пользователей исходя из обозреваемого в данный момент раздела или страницы веб-приложения, помимо учета семантической релевантности ответа запросу. При этом остается потребность в составлении универсальных ответов без учета дополнительных условий, относящихся ко всем страницам веб-приложения, для сокращения используемых объемов хранилищ данных и ускорения обработки поступающих запросов.

В статье описываются статус-кво в области алгоритмов ранжирования, условия систем помощи для применения алгоритмов ранжирования и разработанный алгоритм контекстного ранжирования ответов по относительным метрикам в диалоговых системах помощи пользователям, приводится пример применения алгоритма.

Целью статьи является описание алгоритма ранжирования ответов на запросы пользователей в соответствии с контекстом запроса в диалоговых системах, основанных на списках ответов.

Данный алгоритм рекомендуется применять в диалоговых системах, основанных на списках, в которых релевантность запроса пользователя и ответов из списка основана не только на семантической близости запроса и ответа, но и на релевантности условий применения ответа из списка и условий, в которых пользователь находится в момент запроса. К таким системам, помимо систем помощи пользователям, относятся системы, работающие по принципу «обмен сообщениями как интерфейс» (messaging-as-an-interface). В заключении статьи обозначены перспективные направления дальнейших исследований.

**Ключевые слова:** диалоговые системы, вопросно-ответные системы, обработка естественного языка, искусственный интеллект, ранжирование, обмен сообщениями как интерфейс.

Диалоговые системы активно применяются для оказания информационных услуг во многих сферах, и сейчас часто используется подход «обмен сообщениями как интерфейс» (messaging-as-an-interface), при котором в качестве интерфейса системы выступает чат-бот. Актуальности различных видов диалоговых систем в общем и диалоговых систем, основанных на естественном языке, в частности, посвящены работы [1–3].

Выбор такого подхода может быть обусловлен необходимостью снижения порога для входа пользователя. То есть классические инструменты интерфейсов с множеством элементов управления (кнопки, поля ввода, формы загрузки и т.д.) значительно эффективнее применять в простых системах, например, в системах, состоящих из одного окна интер-

фейса, однако при увеличении сложности системы и необходимости переходить из одного окна к другому в поиске требуемой функции целесообразнее использовать диалоговый режим, когда достаточно указать команду в поле ввода сообщения. Тем более, что при возможности ввода команд в свободной (естественно-языковой) форме не требуются изучение и запоминание большого объема справочной информации, инструкций по системе.

Использование диалоговой системы целесообразно в сложных программах, где велики затраты времени на изучение инструкций. Под сложной программой понимается программный продукт, в котором содержится множество окон или страниц с множеством функций. Так как определение степени сложности системы не является задачей статьи, данный

аспект здесь не рассматривается. О важности эффективности интерфейса человек–машина говорится в работах [4–6].

По методам, используемым для порождения ответов, диалоговые системы можно разделить на следующие типы (в соответствии с работой [7]):

- системы, основанные на правилах, когда ответ на запрос пользователя формируется на основе заранее заданного шаблона или правила;
- системы, основанные на заданном списке ответов, которые ответы на заданные пользователем сообщения выбирают из подготовленного заранее набора;
- системы, основанные на порождающих моделях и генерирующие ответ с помощью некоторой порождающей модели.

Поскольку в данной статье речь идет о диалоговых системах помощи пользователям сложных веб-приложений, использование системы на порождающих моделях не рационально в силу предварительно заданных функций веб-приложения, которые либо можно описать в заранее заданном наборе, либо сформировать правила для их поиска.

В статье рассматривается универсальная система, позволяющая специалистам с низкой квалификацией адаптироваться под любые веб-приложения. Использование систем, основанных на правилах, приводит к затруднениям, так как это более сложный для изучения подход, требующий от специалиста, создающего сценарий обучения и подсказок, знания многих аспектов задания правил. Использование системы, основанной на заданном списке ответов, является более подходящим выбором, так как от специалиста требуется лишь пополнять список, состоящий из вопросов и ответов на естественном языке.

### Системы помощи

В настоящее время получили развитие такие сервисы внедрения и обучения пользователей сложных программных продуктов, как WalkMe, WhatFix, Pendo, Appcues. Они помогают пользователям разобраться в сложном ПО, пошагово указывая на элементы интерфейса, с которыми ему нужно взаимодействовать в ходе исполнения профессиональных обязанностей. О пользе от внедрения подобного ПО говорится в работе [8].

При реализации таких сервисов в формате диалоговых систем возникает сложность с контекстной зависимостью ответа на запрос пользователя от того окна или той страницы приложения, где он сейчас находится. Например, при запросе «Как зарегистрировать аккаунт?» пользователь может иметь в виду регистрацию личного аккаунта в социальной сети, если он находится на главной странице социальной сети, либо регистрацию аккаунта компании, если он уже вошел в личный аккаунт ранее. Таким образом, необходимо ранжировать сценарии помощи в зависимости от релевантности запросу (точность поиска по запросу, соответствие смыслу запроса) и от контекста задаваемого вопроса (расположения пользователя в системе) и выдавать результат с наивысшим рангом. В примере приведен случай, когда необходимо реагировать на адрес страницы, на которой находится пользователь, но, помимо этого, могут обрабатываться и другие параметры: содержание cookie, история переходов, о чем речь идет в работе [9] и других. Также важным моментом, который не рассматривается в данной статье, является реакция на факты, выявленные в ходе диалога с пользователем, о чем упоминается в [10].

### Алгоритм ранжирования

Большое значение алгоритмы ранжирования имеют в поисковых системах, так как ранжирование выдаваемой информации – один из критериев эффективности их работы. Такой же вывод можно сделать из работ [9, 11]. Актуальные алгоритмы ранжирования поисковых систем закрыты, поскольку они являются коммерческой тайной, интеллектуальной собственностью компаний-разработчиков систем. Закрытость алгоритмов ранжирования защищает поисковую систему от мошеннического продвижения сайтов в данной поисковой системе, так как в соответствии с законом Гудхарта или законом Кэмпбелла, приведенными в [12], известность алгоритма оценки приводит к профанации данных метрик.

На текущий момент ранжирование является актуальной проблемой [13, 14]. Как следует из работ [14, 15], сложно выделить алгоритм ранжирования из алгоритма поиска – часто такие алгоритмы описываются в комплексе. Авторы [13] приводят следующие модели поиска и ранжирования: булевая модель, векторно-пространственная модель, вероятностная модель. Они называют недостатки этих моделей и рас-

сравнивают использование коэффициента корреляции, решают задачи повышения быстродействия и снижения требований к аппаратной части систем.

В данной статье затрагиваются вопросы использования относительных метрик для ранжирования ответов в диалоговых системах помощи, то есть недостатки рассмотренных систем остаются актуальными для рассматриваемой предметной области, а имеющиеся алгоритмы их не решают. Эта задача выявлена в ходе разработки автором сервиса онбординга – системы создания обучающих сценариев для новых пользователей веб-приложений.

Исходя из внедряемого посредством системы веб-приложения, набор функций имеет конечное количество. Ответ из набора выбирается в соответствии с метриками релевантности. В связи с использованием множества метрик для определения степени релевантности запроса пользователя вариантам ответов, хранящихся в БД системы, и необходимостью определить единственный наиболее релевантный вариант изначально используется ранжирование самих метрик, чтобы снизить вероятность коллизий релевантностей.

Ранжирование метрик оценки результата по важности или по степени релевантности к рассматриваемой сфере – часто нетривиальная задача из-за сложности подбора метрик оценки самих метрик, но возможно использование не количественных, а качественных показателей (например, какая из сравниваемых метрик более релевантна требованиям системы). Также периодически такие метрики не имеют коррелирующих единиц измерения для других метрик. Для ее решения можно использовать эвристический подход: составить список всех метрик и воспользоваться алгоритмом сортировки (например, сортировки Хоара) для ранжирования списка по относительному показателю степени релевантности метрики требованиям системы. Это применительно на подготовительном этапе системы, когда составляется список метрик релевантности. В соответствии с рангом метрикам выставляются их приоритеты. Приоритеты метрик будут использоваться как множители значений этих метрик сравниваемых результатов – это механизм использования приоритетов метрик в расчетах ранга сравниваемых ответов пользователю.

Так как метрики могут отличаться по порядку значений, необходимо использовать не абсолютные их значения, а относительные:

сравнивать результаты по каждой метрике между собой и учитывать результат этого сравнения. Тогда значения всех метрик, даже не имеющих количественных выражений, будут соизмеримы и влияние на разницу между метриками будет иметь только выбранный приоритет метрики.

В качестве значений приоритетов и метрик результатов можно использовать дробные числа – это позволит расставлять значения сразу в момент сравнения метрик или результатов по этим метрикам, так как дробные значения могут дробиться до бесконечности, а выбрать целое число между целыми числами периодически бывает невозможно. То есть, вместо того чтобы сначала отсортировать значения, а потом присвоить им численные представления ранга, можно назначить начальное значение ранга первому выбранному элементу и присваивать значения остальным относительно первого выбранного элемента. Для приоритетов это может быть число 0,5. В качестве изначального значения первого сравниваемого результата предлагается брать значение приоритета метрики: это приведет к лавинообразному росту разницы между значениями разных метрик и снизит вероятность коллизий.

Итоговая форма представления сравниваемых ответов (результатов) и их метрик приведена в таблице 1.

Таблица 1

**Представление сравниваемых результатов и их метрик**

Table 1

**The compared results' presentation and their metrics**

Метрика	Приоритет	Результат			
		1	2	...	M
1	$p[1]$	$v[1][1]$	$v[2][1]$	...	$v[M][1]$
2	$p[2]$	$v[1][2]$	$v[2][2]$	...	$v[M][2]$
...	...	...	...	...	...
$N$	$p[N]$	$v[1][N]$	$v[2][N]$	...	$v[M][N]$
Ранг результата	–	$r[1]$	$r[2]$	...	$r[M]$

Для повышения порядковой разницы значения метрик каждого результата перемножаются с приоритетами соответствующих метрик, это позволяет задать приоритет значений разных метрик для каждого результата. Далее значения метрик каждого результата складываются. Таким образом, ранг релевантности результата  $i$  или  $r[i]$  равен  $r[i] = \sum_{j=1}^N v[i][j] \times p[j]$ .

где  $v[i][j]$  – значение метрики  $j$  результата  $i$ ;  $p[j]$  – значение приоритета метрики  $j$ .

При подобной работе гипотетически возможны коллизии. Для их устранения также предлагается алгоритмическое решение. Если коллизия возникла на максимальном значении ранга, все результаты с максимальным рангом поочередно сравниваются по метрикам от самой приоритетной метрики к наименее приоритетной. Результат с первым найденным высшим значением метрики считается результатом с более высоким приоритетом. Гипотетически возможна ситуация полного совпадения, тогда предлагается брать результат, значение которого стоит выше при сортировке по алфавиту. Либо необходимо исключить возможность равенства значений метрик для результатов, то есть при сравнении метрик допускать только условия «более релевантно» или «менее релевантно» без возможности равенства релевантностей. При таком подходе коллизия будет устраняться за один шаг.

#### Пример использования алгоритма

Для примера выбран запрос пользователя на создание аккаунта компании в социальной сети ВКонтакте: «Как создать аккаунт». Данный запрос сделан, когда пользователь находился на одной из внутренних страниц социальной сети, то есть уже вошел в свой личный аккаунт и, вероятнее всего, имел в виду создание корпоративного аккаунта. Это упрощенный пример для иллюстрации работы алгоритма.

В качестве метрик простейшего представления системы помощи пользователям выступают релевантность запроса пользователя возможным запросам, хранящимся в БД, и релевантность веб-страницы, на которой находится пользователь. Метрика релевантности страницы, на которой находится пользователь, имеет больший приоритет, так как важнее, на какой из страниц осуществляется подобный запрос, нежели качество распознавания запроса. В соответствии с представленными условиями составлена таблица 2.

Полученные результаты показывают, что у сценария «Создать аккаунт компании» большее значение ранга, значит, необходимо запустить данный сценарий помощи пользователю.

Таблица 2

#### Пример представления сравнимых результатов и их метрик

Table 2

#### The presentation example of the compared results and their metrics

Метрика	Приоритет	Создать аккаунт компании	Создать личный аккаунт
Релевантность страницы	0,6	0,7	0,6
Релевантность запроса	0,5	0,4	0,5
Ранг результата	–	0,62	0,61

#### Заключение

В статье представлен обзор разработанного алгоритма для ранжирования вариантов ответов на запрос пользователя. Алгоритм ранжирования применим не только в системах помощи пользователям, но и в других системах сравнения объектов по множеству не коррелирующих между собой метрик, при этом метрики могут не иметь численного представления, то есть быть качественными, а не количественными. Данный алгоритм решает поставленную задачу.

При дальнейшей работе можно рассмотреть влияние фактов, выявленных в ходе диалога с пользователем, на релевантность ответов запросам, о чем упоминается в [10]. Возможны оптимизация работы с коллизиями, оптимизация процесса присвоения значений приоритетов и значений метрик, сравнения различных алгоритмов присвоения значений метрик результатам, влияние выделения терминологии на качество ранжирования, эффективность чего показана в работе [11]. Следует рассмотреть автоматизацию построения ранжирующей функции исходя из множества параметров, что, как следует из [9], является актуальным. Также возможны исследование применения данного алгоритма ранжирования в других задачах и сравнение с другими алгоритмами ранжирования в сферах, где такое сравнение осуществимо. Важным направлением является реализация алгоритма с использованием современных достижений техники и оценкой влияния, например, бессерверной архитектуры на эффективность работы, о чем говорится в [16].

**Литература**

1. Gafurov N.R., Bessmertny I.A., Platonov A.V., Poleshchuk E.A., Vasiliev A.V. Named entity recognition through bidirectional lstm in natural language texts obtained through audio interfaces. Proc. 12th IEEE Intern. Conf. AICT, Almaty, Kazakhstan, 2018, pp. 74–77.
2. Бессмертный И.А., Посевкин Р.В. Естественно-языковой пользовательский интерфейс диалоговой системы // Программные продукты и системы. 2016. № 3. С. 5–9. DOI: 10.15827/0236-235X.115.005-009.
3. Крайванова В.А. Модель естественно-языкового интерфейса для систем управления сложными техническими объектами и оценка эффективности алгоритмов на ее основе // Управление большими системами. Вып. 26. 2009. С. 158–178.
4. Kujala S., Roto V., Väänänen-Vainio-Mattila K., Karapanos E., Sinnelä A. UX Curve: a method for evaluating long-term user experience. *Interacting with Computers*, 2011, vol. 23, no. 5, pp. 473–483. DOI: 10.1016/j.intcom.2011.06.005.
5. Marcus A. Design, user experience, and usability: Design discourse. Proc. 4th Intern. Conf. DUXU, 2015, 672 p.
6. Ахметов К. Взаимодействие человека и компьютера: тенденции, исследования, будущее // Форсайт. 2013. Т. 7. № 2. С. 58–68.
7. Чистова Е.В., Шелманов А.О., Смирнов И.В. Применение глубокого обучения к моделированию диалога на естественном языке // Тр. ИСА РАН. 2019. Т. 69. № 1. С. 105–115. DOI: 10.14357/20790279190110.
8. Lepofsky A. Digital Adoption Platforms: a new breed of software helps improve user adoption of new technologies. Constellation Research, 2017. URL: <https://blog.walkme.com/wp-content/uploads/2017/05/20170512-CR-A-LEPOFSKY-DIGITAL-ADOPTION-PLATFORMS-FINAL-1-1.pdf> (дата обращения: 15.08.2019).
9. Николенко С.И., Фишков А.А. Обзор моделей поведения пользователей для задачи ранжирования результатов поиска // Тр. СПИИРАН. 2012. Вып. 22. С. 139–175.
10. Bessmertny I. On constructing intellectual systems in ternary logic. *Programming and Computer Software*, 2014, vol. 40, no. 1, pp. 43–46.
11. Заикин Д.А. Подход к ранжированию результатов для терминологического поиска на основе меры близости строк // Ученые записки Казан. ун-та. 2014. Т. 156. С. 12–21.
12. Форум: Прикладная наукометрия // Антропологический форум. 2019. № 40. С. 11–84. URL: <http://anthropologie.kunstkamera.ru/files/pdf/040/forum.pdf> (дата обращения: 15.08.2019). DOI: 10.31250/1815-8870-2019-15-40-11-84.
13. Кучукова Н.Н., Вершков Н.А. Математическая модель подсистемы поиска и ранжирования документов в информационно-поисковых системах // Научные ведомости БелГУ. 2018. № 1. С. 176–183.
14. Брицов Р.А. Повышение качества информационного поиска за счет совершенствования ранжирования и использования особенностей поведения пользователей // Т-Comm: Телекоммуникации и транспорт. 2016. Т. 10. № 2. С. 63–66.
15. Brin S., Page L. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 1998, vol. 30, iss. 1–7, pp. 107–117.
16. Yan M., Castro P., Cheng P., Ishakian V. Building a chatbot with serverless computing. Proc. 1st Intern. Workshop on Mashups of Things and APIs MOTA, 2016, pp. 1–4.

**Response contextual ranking on relative metrics in dialogue assistance systems for users**

*N.R. Gafurov*<sup>1</sup>, *Postgraduate Student, nail@gafurov.ru*

<sup>1</sup>*The National Research University of Information Technologies, Mechanics and Optics, St. Petersburg, 197101, Russian Federation*

**Abstract.** The subject of this paper is dialogue systems for assisting users of web applications or onboarding systems.

The paper contains dialog systems in general, the probable reasons for their current updating, the conditions for their feasibility, their types, and justifies the choice of the type for the described system.

We have studied the help systems that make it necessary to answer users' questions based on the currently viewed section or a web application page, in addition to taking into account the semantic relevance of the response to the query.

At the same time there remains the need to give universal answers without taking into account additional conditions that apply to all pages of a web application to reduce the amount of data storage used and speed up the processing of incoming requests.

The paper describes the status quo in the ranking algorithm field, the system conditions of assisting users for applying ranking algorithms and describes the developed algorithm for response contextual ranking by relative metrics in dialogue systems for assisting users provides the algorithm application example.

The paper's purpose is to describe the algorithm for ranking responses to user requests in accordance with the request context in dialogue systems based on answer lists.

This algorithm is for use in dialogue systems based on lists in which the user's request relevance and responses from the list are based not only on the request semantic proximity and response but also on the condition relevance for applying the response from the list and the conditions in which the user is at the time of the request. Such systems in addition to user assistance systems, include systems operating on the "messaging-as-an-interface" principle. The paper also gives promising areas for further research.

**Keywords:** dialogue systems, question-answer systems, natural language processing, artificial intelligence, ranking, messaging as an interface.

### References

1. Gafurov N.R., Bessmertny I.A., Platonov A.V., Poleshchuk E.A., Vasiliev A.V. Named entity recognition through bidirectional lstm in natural language texts obtained through audio interfaces. *Proc. 12th IEEE Intern. Conf. AICT*, Almaty, Kazakhstan, 2018, pp. 74–77.
2. Bessmertny I.A., Posevkin R.V. Natural language user interface of a dialogue system. *Software & Systems*, 2016, no. 3, pp. 5–9. DOI: 10.15827/0236-235X.115.005-009 (in Russ.).
3. Krayvanova V.A. Model of natural-language interface for complex technical objects control systems and efficiency evaluation of model-based algorithms. *Large-Scale Systems Control*, 2009, iss. 26, pp. 158–178 (in Russ.).
4. Kujala S., Roto V., Väänänen-Vainio-Mattila K., Karapanos E., Sinnelää A. UX Curve: a method for evaluating long-term user experience. *Interacting With Computers*, 2011, vol. 23, no. 5, pp. 473–483. DOI: 10.1016/j.intcom.2011.06.005.
5. Marcus A. Design, user experience, and usability: Design discourse. *Proc. 4th Intern. Conf. DUXU*, 2015, 672 p.
6. Akhmetov K. Human-computer interaction: Trends, research, future. *Foresight and STI Governance*, 2013, vol. 7, no. 2, pp. 58–68 (in Russ.).
7. Chistova E.V., Shelmanov A.O., Smirnov I.V. Natural language dialogue modelling with deep learning. *Proc. of the Institute for Systems Analysis Rus. Academy of Sc.*, 2019, vol. 69, no. 1, pp. 105–115. DOI: 10.14357/20790279190110 (in Russ.).
8. Lepofsky Alan. Digital adoption platforms: A new breed of software helps improve user adoption of new technologies. *Constellation Research*, 2017. Available at: <https://blog.walkme.com/wp-content/uploads/2017/05/20170512-CR-A-LEPOFSKY-DIGITAL-ADOPTION-PLATFORMS-FINAL-1-1.pdf>. (accessed August 15, 2019).
9. Nikolenko S.I., Fishkov A.A. A survey of user behaviour models for search results ranking. *Proc. SPIIRAN*, 2012, iss. 22, pp. 139–175 (in Russ.).
10. Bessmertny I. On constructing intellectual systems in ternary logic. *Programming and Computer Software*, 2014, vol. 40, no. 1, pp. 43–46.
11. Zaikin D.A. A string proximity-based approach to result ranking for terminological search. *Sc. Research in KFU*, 2014, vol. 156, pp. 12–21 (in Russ.).
12. *Forum for Anthropology and Culture*. 2019, no. 40, pp. 11–84. Available at: <http://anthropologie.kunstkamera.ru/files/pdf/040/forum.pdf> (accessed August 15, 2019). DOI: 10.31250/1815-8870-2019-15-40-11-84.
13. Kuchukova N.N., Vershkov N.A. Mathematical model of subsystem of document retrieval and ranking in information-retrieval systems. *Belgorod State University Sc. Bull.*, 2018, no. 1, pp. 176–183. DOI: 10.18413/2411-3808-2018-45-1-176-183 (in Russ.).

14. Britsov R.A. Improvement of quality of the information search based on ranking rationalization and users' behavioristic characteristics. *T-Comm*, 2016, vol. 10, no. 2, pp. 63–66 (in Russ.).

15. Brin S., Page L. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 1998, vol. 30, iss. 1–7, pp. 107–117.

16. Yan M., Castro P., Cheng P., Ishakian V. Building a chatbot with serverless computing. *Proc. 1st Intern. Workshop on Mashups of Things and APIs MOTA*, 2016, pp. 1–4.

#### Для цитирования

Гафуров Н.Р. Контекстное ранжирование ответов по относительным метрикам в диалоговых системах помощи пользователям // Программные продукты и системы. 2020. Т. 33. № 2. С. 276–282. DOI: 10.15827/0236-235X.130.276-282.

#### For citation

Gafurov N.R. Response contextual ranking on relative metrics in dialogue assistance systems for users. *Software & Systems*, 2020, vol. 33, no. 2, pp. 276–282 (in Russ.). DOI: 10.15827/0236-235X.130.276-282.