

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРИНЯТИЕ РЕШЕНИЙ
ARTIFICIAL INTELLIGENCE AND DECISION MAKING

УДК 004.03

DOI: 10.18413/2518-1092-2022-8-3-0-4

Бузов П.А.¹
Игитян Е.В.²
Польщиков К.А.²
Лебедева А.С.²

ОЦЕНИВАНИЕ ВЕРОЯТНОСТНО-ВРЕМЕННЫХ
ХАРАКТЕРИСТИК ЧЕЛОВЕКО-МАШИННОГО ДИАЛОГА

¹) АО «СОФТКОННЕКТ», ул. Рабочая, 14, офис 1 Белгород, 308013, Россия
²) Белгородский государственный национальный исследовательский университет,
ул. Победы, 85, Белгород, 308015, Россия

e-mail: medvedeva_e@bsu.edu.ru

Аннотация

Работа посвящена оцениванию вероятностно-временных характеристик человеко-машинного диалога. Авторами проведен ряд экспериментов по исследованию характеристик диалоговой системы «Алиса» и оцениванию показателей, характеризующих особенности её пользователей. Результат показал, что отклонения характеристик человеко-машинных диалогов, полученных в результате экспериментов, от характеристик, вычисленных в результате моделирования, не превышают $\pm 3,8$ % при доверительной вероятности 0,95. В частности, авторами разработана модель человеко-машинного диалога, позволяющая учитывать характеристики диалоговой системы.

Ключевые слова: диалоговые системы; эффективность чат-бот-приложений; человеко-машинный диалог

Для цитирования: Бузов П.А., Игитян Е.В., Польщиков К.А., Лебедева А.С. Оценивание вероятностно-временных характеристик человеко-машинного диалога // Научный результат. Информационные технологии. – Т.8, №3, 2023. С. 27-33. DOI: 10.18413/2518-1092-2022-8-3-0-4

Buzov P.A.¹
Igityan E.V.²
Polshchikov K.A.²
Lebedeva A.S.²

ESTIMATION OF PROBABILISTIC-TEMPORAL
CHARACTERISTICS OF HUMAN-MACHINE DIALOGUE

¹) JSC "SOFTCONNECT", 14 Rabochaya str., office 1 Belgorod, 308013, Russia
²) Belgorod State National Research University, 85 Pobedy Str., Belgorod, 308015, Russia

e-mail: medvedeva_e@bsu.edu.ru

Abstract

The paper is devoted to the evaluation of the probabilistic-temporal characteristics of the human-machine dialogue. The authors conducted a number of experiments to study the characteristics of the Alice dialog system and evaluate the indicators characterizing the characteristics of its users. The result showed that the deviations of the characteristics of human-machine dialogues obtained as a result of experiments from the characteristics calculated as a result of modeling do not exceed $\pm 3.8\%$ with a confidence probability of 0.95. In particular, the authors have developed a human-machine dialogue model that allows taking into account the characteristics of the dialog system.

Keywords: dialog systems; efficiency of chatbot applications; human-machine dialogue

For citation: Buzov P.A., Igityan E.V., Polshchikov K.A., Lebedeva A.S. Estimation of probabilistic-temporal characteristics of human-machine dialogue // Research result. Information technologies. – Т.8, №3, 2023. – P. 27-33. DOI: 10.18413/2518-1092-2022-8-3-0-4

ВВЕДЕНИЕ

Одним из важнейших направлений автоматической обработки естественно-языковых данных является разработка и совершенствование интеллектуальных диалоговых систем и их упрощенных версий – чат-ботов. Эти системы стали все чаще применяться в коммерческих проектах, где они используются в общении с клиентами для помощи в покупке товаров, технической поддержки, навигации по сайтам и т.д. Диалоговые системы используются в качестве интеллектуальных модулей общения социальных роботов, которые ухаживают за больными, престарелыми людьми.

Задача диалоговых систем – не только продемонстрировать пользователю свои коммуникативные способности, но и предоставить максимально точный ответ на вопрос, заданный на естественном языке [5]. Диалоговые системы, основанные на использовании программных средств обработки естественно-языковых данных и предназначенные для выдачи ответов на заданные пользователями вопросы, будем именовать вопросно-ответными системами. Современные вопросно-ответные системы обладают модулями контентной аналитики, позволяющими собирать и упорядочивать информацию, а также использовать машинное обучение на основе нейронных сетей. Известными примерами таких систем являются вопросно-ответные сервисы «ChatGPT» (OpenAI), «Watson» (IBM), виртуальные ассистенты «Алиса» (Яндекс), «Siri» (Apple), «Google Assistant», «Amazon Alexa», «Cortana» (Microsoft). Специалисты утверждают, что существует потребность в адекватных средствах оценивания характеристик диалоговых систем. Недостает общепризнанных методов, моделей, метрик, критериев и количественных показателей, на основе которых можно было бы обеспечить обоснованность принимаемых решений о качестве функционирования диалоговых систем, объективно выявить их недостатки и определить направления дальнейшего совершенствования.

Существующие инструментарии оценки характеристик диалоговых систем позволяют оценить обобщенный показатель выполнения диалоговой системой вопросно-ответных функций без привязки к особенностям тех или иных пользователей или их групп [2]. При этом в таких средствах оценки параметры нечеткого вывода настроены на «среднего пользователя» и не учитывают такие человеческие особенности, как, например, терпеливость в задавании вопросов или сообразительность при получении недостаточного точного и полного ответа и т.п. Чтобы обеспечить учет подобных пользовательских особенностей потребуется многократное усложнение предложенных алгоритмов оценивания, основанных на применении нечеткого вывода и нейросетевой настройки, т.к. для каждой группы пользователей будет необходимо:

- вычислить частные показатели точности, лаконичности и полноты ответов;
- сформировать отдельные обучающие выборки;
- создать отдельные системы нечеткого вывода;
- настроить параметры каждой системы нечеткого вывода с помощью нейросетевого обучения;
- оценить обобщенный показатель выполнения диалоговой системой вопросно-ответных функций, реализовав процедуры нечеткого вывода.

Такое усложнение процесса вычисления характеристик диалоговой системы неприемлемо, поэтому для оценивания соответствия характеристик вопросно-ответной системы особенностям её пользователей требуется модель человеко-машинного диалога, результаты разработки которой представлены ниже.

РАЗРАБОТКА МОДЕЛИ ЧЕЛОВЕКО-МАШИННОГО ДИАЛОГА

Пусть у пользователя имеется потребность в получении определенной информации. Для удовлетворения этой потребности пользователь вступает в человеко-машинный диалог. Следовательно, целью этого диалога является удовлетворение определенной информационной потребности пользователя. На рисунке 1 представлен вероятностно-временной граф, моделирующий процесс человеко-машинного диалога.

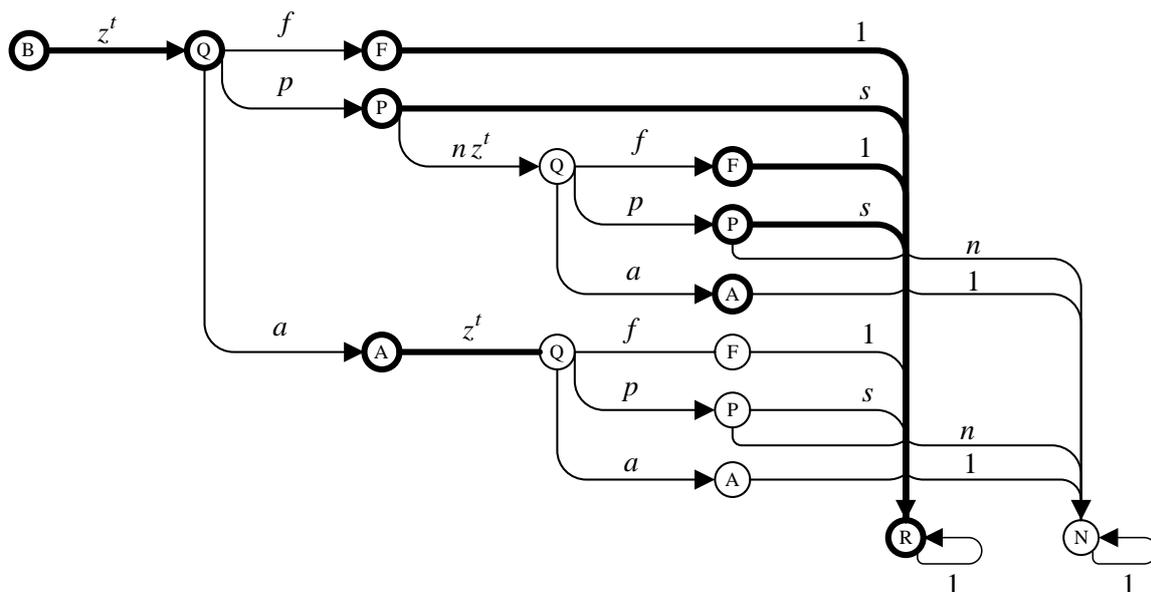


Рис. 1. Вероятностно-временной граф, моделирующий процесс человеко-машинного диалога
Fig. 1. Probabilistic-time graph modeling the process of human-machine dialogue

Начальное состояние моделируемого процесса обозначено «В». В ходе диалога человеком задается вопрос, при этом процесс моделирования переходит в состояние «Q». Далее на заданный вопрос система выдает ответ. Если полученный пользователем ответ в полной мере содержит требуемую информацию (такой вариант соответствует переходу из состояния «Q» в состояние «F»), то считается, что цель диалога достигнута, и моделируемый процесс переходит из состояния «F» в терминальное состояние «R». Если в полученном ответе практически отсутствует требуемая информация (переход из состояния «Q» в состояние «A»), пользователь может прекратить диалог без достижения цели (переход из состояния «P» в терминальное состояние «N»), посчитав систему бесполезной для удовлетворения его информационной потребности, а может задать системе уточняющий вопрос (переход из состояния «P» в состояние «Q»).

В случае получения ответа, в котором частично содержится требуемая информация, моделируемый процесс переходит из состояния «Q» в состояние «P». При этом возможны следующие варианты:

- 1) пользователь удовлетворяется полученным ответом, цель диалога достигается, а моделируемый процесс переходит из состояния «P» в терминальное состояние «R»;
- 2) пользователь не удовлетворяется полученным ответом и задает уточняющий вопрос, моделируемый процесс переходит из состояния «P» в состояние «Q»;
- 3) пользователь не удовлетворяется полученным ответом и прекращает задавать вопросы, диалог заканчивается, не достигнув цели, что соответствует переходу из состояния «P» в терминальное состояние «N».

В результате выполнения конечного числа указанных выше переходов диалог завершается, т.е. моделируемый процесс оказывается в терминальном состоянии «R» или в терминальном состоянии «N». Граф, изображенный на рисунке 1, построен для случая, при котором число задаваемых пользователем уточняющих вопросов, ограничивается величиной $m = 1$. Переход по стрелке из какой-либо вершины графа в соседнюю вершину моделируется с помощью некоторой функции:

$$\Omega(z) = \omega \cdot z^\tau, \quad (1)$$

где ω – вероятность осуществления перехода; τ – время, в течение которого осуществляется переход; z – вспомогательный параметр, используемый в теории производящих функций.

Переход из состояния «N» в состояние «Q» моделируется функцией z^t , где t – средняя продолжительность по времени формулировки пользователем вопроса и выдачи системой ответа.

Переходу из состояния «Q» в состояние «F» соответствует величина f – вероятность того, что выданный системой ответ в полной мере содержит требуемую пользователю информацию. Переходу из состояния «Q» в состояние «P» соответствует величина p – вероятность выдачи системой ответа, в котором частично содержится требуемая информация. Из состояния «Q» возможен также переход в состояние «A». Этому событию соответствует величина a – вероятность отсутствия в выданном системой ответе информации, требуемой пользователю. По условию нормировки вероятность a может быть вычислена с помощью выражения:

$$a = 1 - f - p.$$

Из состояния «F» в терминальное состояние «R» осуществляется переход с вероятностью, равной 1. Переходу из состояния «P» в состояние «R» соответствует величина s – вероятность удовлетворенности пользователя полученным ответом, в котором частично содержится требуемая информация. Переход из состояния «P» в состояние «Q» моделируется функцией nz^t , где n – вероятность неудовлетворенности пользователя полученным ответом, в котором частично содержится требуемая информация. По условию нормировки вероятностей величину n можно вычислить с помощью выражения:

$$n = 1 - s.$$

Переход из состояния «A» в состояние «Q» моделируется функцией z^t . Переходу из состояния «P» в состояние «N» соответствует вероятность n . Наконец, переход из состояния «A» в терминальное состояние «N» осуществляется с вероятностью, равной 1. В результате эквивалентных преобразований графа, получены выражения для вычисления $\Omega_{BR}^1(z)$ – функции перехода из состояния «B» в состояние «R» при $m=1$, $\Omega_{BN}^1(z)$ – функции перехода из состояния «B» в состояние «N» при $m=1$, $\Omega_{BE}^1(z)$ – функции перехода из состояния «B» в состояние «E» при $m=1$. Удалось вывести соответствующие выражения для $m=2$, $m=3$ и $m=4$, выявить в них обобщающие закономерности, на основе которых получены формулы для вычисления функций $\Omega_{BR}^m(z)$, $\Omega_{BN}^m(z)$ и $\Omega_{BE}^m(z)$ при любом натуральном m :

$$\Omega_{BR}^m(z) = (f + ps)z^t \sum_{\alpha=0}^m \sum_{\beta=0}^{m-\alpha} (pnz^t)^\alpha (\alpha z^t)^\beta \frac{(\alpha + \beta)!}{\alpha! \beta!}, \quad (2)$$

$$\Omega_{BN}^m(z) = (pn + a)z^{(m+1)t} \sum_{\alpha=0}^m (pn)^\alpha a^{(m-\alpha)} \frac{m!}{\alpha! (m-\alpha)!}, \quad (3)$$

$$\Omega_{BE}^m(z) = \Omega_{BR}^m(z) + \Omega_{BN}^m(z). \quad (4)$$

Для вычисления вероятности достижения цели диалога можно использовать выражение:

$$r_m = \Omega_{BR}^m(z) \Big|_{z=1}. \quad (5)$$

Средняя продолжительность человеко-машинного диалога может быть вычислена по формуле:

$$u_m = \frac{d}{dz} \Omega_{BE}^m(z) \Big|_{z=1}. \quad (6)$$

РЕЗУЛЬТАТЫ РАБОТЫ

На основе применения разработанной модели проведены экспериментальные исследования по оцениванию вероятностно-временных характеристик человеко-машинного диалога. Графики зависимости величин r_m и u_m от значений величины m при $f=0,3$ и $p=0,4$ представлены на рисунках 2 и 3.

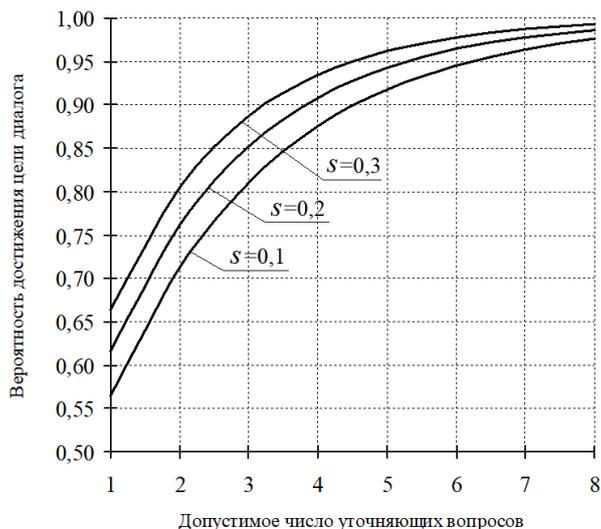


Рис. 2. Графики зависимости величины r_m от значений величины m при $f = 0,3$ и $p = 0,4$

Fig. 2. Graphs of the dependence of the magnitude on the values r_m of the magnitude m at $f = 0,3$ and $p = 0,4$

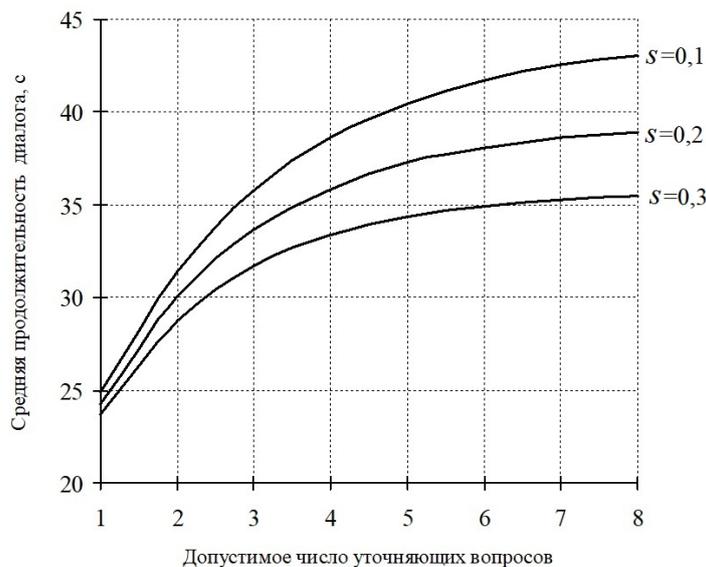


Рис. 3. Графики зависимости величины u_m от значений величины m при $f = 0,3$ и $p = 0,4$

Fig. 3. Graphs of the dependence of the magnitude on the values u_m of the magnitude m at $f = 0,3$ and $p = 0,4$

Допустим заданы требования к характеристикам диалоговой системы, например, вероятность достижения цели человеко-машинного диалога не должна быть ниже 0,9, а средняя продолжительность диалога при этом не должна превышать 40 с. Тогда анализ представленных на рисунках 2 и 3 графиков показывает, что диалоговая система с параметрами $f = 0,3$ и $p = 0,4$ рекомендуется тем пользователям, которые, в случае необходимости, готовы задать хотя бы 4 уточняющих вопроса, а также с вероятностью не ниже 0,2 будут удовлетворены полученными ответами, в которых лишь частично содержится требуемая информация.

Проведены многочисленные эксперименты по исследованию характеристик диалоговой системы «Алиса» и оцениванию показателей, характеризующих особенности её пользователей. В ходе экспериментов проводились диалоги пользователей с этой системой, фиксировались результаты, характеризующие содержание требуемой информации в выдаваемых системой ответах и удовлетворенность ими пользователей. В экспериментальных диалогах приняли участие 134 пользователя, каждый из которых провел с системой по 20 диалогов. Выполнена обработка экспериментальных данных, проведенных с участием всех пользователей, а затем получены результаты вычислений на основе разработанной модели.

Проведен статистический анализ отклонений результатов экспериментов от результатов моделирования. Анализ показал, что отклонения характеристик человеко-машинных диалогов, полученных в результате экспериментов, от характеристик, вычисленных в результате моделирования, не превышают $\pm 3,8\%$ при доверительной вероятности 0,95. Высокая сходимость результатов экспериментальных исследований и моделирования свидетельствует об адекватности разработанной модели человеко-машинного диалога.

ЗАКЛЮЧЕНИЕ

Модель человеко-машинного диалога учитывает характеристики диалоговой системы, в числе которых следующие величины: вероятность выдачи системой ответа, в котором частично содержится требуемая информация; вероятность того, что выданный системой ответ в полной мере содержит требуемую пользователю информацию; вероятность отсутствия в выданном системой ответе информации, требуемой пользователю. Кроме того, модель учитывает значения параметров, характеризующих особенности пользователя: вероятность удовлетворенности пользователя полученным ответом, в котором частично содержится требуемая информация; вероятность неудовлетворенности пользователя полученным ответом, в котором частично содержится требуемая информация; максимальное количество уточняющих вопросов, задаваемых пользователем в диалоге.

Список литературы

1. Игитян Е.В., К.А. Польщиков, А.Н. Немцев *Оценивание вероятностно-временных характеристик человеко-машинного диалога на естественном языке. Экономика. Информатика. – 2023. – Т. 50. – № 1. – С. 162–172.*
2. Махди Т.Н., Игитян Е.В., Польщиков К.А., Корсунов Н.И. *Оценивание эффективности функционирования диалоговой системы на основе применения нечеткого вывода с нейросетевой настройкой. Экономика. Информатика. – 2022. – Т. 49. – № 2. – С. 356–374.*
3. Польщиков К.А., Польщикова О.Н., Игитян Е.В., Балакшин М.С. *Алгоритм поддержки принятия решений по выбору средств обработки больших массивов естественно-языковых данных. Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. – 2019. – Т. 46. – № 3. – С. 553-562.*
4. Польщиков К.А., Лазарев С.А., Константинов И.С., Польщикова О.Н., Свойкина Л.Ф., Игитян Е.В., Балакшин М.С. *Модель для оценки эффективности выполнения робототехнической системой коммуникативных функций. СТИН. – 2020. – № 6. – С. 4–7.*
5. Агузумцян Р.В., Великанова А.С., Польщиков К.А., Игитян Е.В., Лихошерстов Р.В. *О применении интеллектуальных технологий обработки естественного языка и средств виртуальной реальности для поддержки принятия решений при подборе исполнителей проектов. Экономика. Информатика. – 2021. – № 48 (2). – С. 392-404.*

References

1. Igityan E.V., Polshchikov K.A., Nemtsev A.N. *Estimation of probabilistic-temporal characteristics of human-machine dialogue in natural language. Economics. Information technologies. – 2023. – Vol. 50. – No. 1. – pp. 162-172.*
2. Mahdi T.N., Igityan E.V., Polshchikov K.A., Korsunov N.I. *Evaluation of the effectiveness of the functioning of a dialog system based on the use of fuzzy inference with neural network setup. Economics. Information technologies. – 2022. – Vol. 49. – No. 2. – pp. 356-374.*

3. Polshchikov K.A., Polshchikova O.N., Igityan E.V., Balakshin M.S. Algorithm of decision support for the choice of means of processing large arrays of natural language data. Scientific Bulletin of Belgorod State University. Series: Economics. Computer science. – 2019. – Vol. 46. – No. 3. – pp. 553-562.

4. Polshchikov K.A., Lazarev S.A., Konstantinov I.S., Polshchikova O.N., Svoikina L.F., Igityan E.V., Balakshin M.S. A model for evaluating the effectiveness of performing communicative functions by a robotic system. STIN. – 2020. – No. 6. – pp. 4-7.

5. Aguzumtsyan R.V., Velikanova A.S., Polshchikov K.A., Igityan E.V., Likhosherstov R.V. On the application of intelligent natural language processing technologies and virtual reality tools to support decision-making in the selection of project performers // Economics. Information technologies. – 2021. – № 48 (2). – Pp. 392-404.

Бузов Павел Андреевич, генеральный директор АО «СОФТКОННЕКТ»

Игитян Елена Владимировна, старший преподаватель кафедры информационных и робототехнических систем, институт инженерных и цифровых технологий

Польщиков Константин Александрович, доцент кафедры информационных и робототехнических систем программного обеспечения вычислительной техники и автоматизированных систем, институт инженерных и цифровых технологий

Лебедева Анна Сергеевна, аспирант 3 курса, направления «Информатика и вычислительная техника», кафедра прикладной информатики и информационных технологий, институт инженерных и цифровых технологий

Buzov Pavel Andreevich, General Director of SOFTCONNECT JSC

Igityan Elena Vladimirovna, Senior Lecturer, Department of Information and Robotic Systems, Institute of Engineering and Digital Technologies

Polshchikov Konstantin Alexandrovich, Associate Professor of the Department of Information and Robotic Systems of Computer Engineering and Automated Systems, Institute of Engineering and Digital Technologies

Lebedeva Anna Sergeevna, 3rd year postgraduate student, Computer Science and Computer Engineering, Department of Applied Computer Science and Information Technology, Institute of Engineering and Digital Technologies