

УДК 004.8
DOI 10.52575/2687-0932-2023-50-1-162-172

Оценивание вероятностно-временных характеристик человеко-машинного диалога на естественном языке

Игитян Е.В., Польщиков К.А., Немцев А.Н.

Белгородский государственный национальный исследовательский университет,
Россия, 308015, г. Белгород, ул. Победы, д. 85
E-mail: medvedeva_e@bsu.edu.ru, polshchikov@bsu.edu.ru, nemzev@bsu.edu.ru

Аннотация. Представлены результаты исследования, направленные на совершенствование процесса оценивания характеристик человеко-машинного диалога, осуществляемого на естественном языке. Обоснована актуальность разработки инструментария для оценивания характеристик вопросно-ответной системы с точки зрения достижения цели человеко-машинного диалога, состоящей в удовлетворении конкретной информационной потребности пользователя. Представлена модель процесса человеко-машинного диалога на основе применения математического аппарата вероятностно-временных графов. Разработанная модель ориентирована на вычисление вероятности достижения цели диалога и его средней продолжительности с учетом характеристик вопросно-ответной системы и значений показателей, характеризующих особенности пользователя. Показано, что применение модели позволяет обосновать выбор диалоговых систем с конкретными характеристиками и рекомендовать их тем или иным группам пользователей для удовлетворения информационных потребностей.

Ключевые слова: обработка естественного языка, диалоговая система, вопросно-ответная система, оценивание вероятностно-временных характеристик, человеко-машинный диалог

Для цитирования: Игитян Е.В., Польщиков К.А., Немцев А.Н. 2023. Оценивание вероятностно-временных характеристик человеко-машинного диалога на естественном языке. Экономика. Информатика, 50(1): 162–172. DOI 10.52575/2687-0932-2023-50-1-162-172

Estimation of Probabilistic and Time Characteristics of Human-Machine Dialogue in Natural Language

Elena V. Igityan, Konstantin A. Polshchikov, Alexander N. Nemtsev

Belgorod National Research University,
85 Pobedy St., Belgorod, 308015, Russia
E-mail: medvedeva_e@bsu.edu.ru, polshchikov@bsu.edu.ru, nemzev@bsu.edu.ru

Abstract. The results of a study aimed at improving the process of evaluating the characteristics of a human-machine dialogue carried out in natural language are presented. The relevance of the development of tools for estimating the probabilistic-temporal characteristics of a question-answer system from the point of view of achieving the goal of a human-machine dialogue, which is to satisfy a specific information need of the user, is substantiated. A model of the human-machine dialogue process based on the use of the mathematical apparatus of probabilistic-time graphs is presented. The developed model is focused on calculating the probability of achieving the goal of the dialogue and its average duration, taking into account the characteristics of the question-answer system and the values of indicators characterizing the user's characteristics. It is shown that the application of the model makes it possible to justify the choice of dialogue systems with specific characteristics and recommend them to certain user groups to meet information needs.

Keywords: natural language processing, dialogue system, question-answer system, estimation of probabilistic-temporal characteristics, human-machine dialogue

For citation: Igityan E.V., Polshchikov K.A., Nemtsev A.N. 2023. Estimation of Probabilistic and Time Characteristics of Human-Machine Dialogue in Natural Language. Economics. Information technologies, 50(1): 162–172 (in Russian). DOI 10.52575/2687-0932-2023-50-1-162-172

Введение

Современные интеллектуальные технологии обработки естественного языка активно применяются для решения различных прикладных задач [Агузумцян и др., 2021; Velikanova et al., 2021; Polshchikov et al., 2022]. На основе таких технологий функционируют диалоговые системы, которые могут быть ориентированы на выполнение конкретных голосовых команд, поддержание общения или работу в вопросно-ответном режиме [Shah et al., 2016; Wen et al., 2017; Deriu et al., 2021]. Анализ функционирования знание-ориентированных диалоговых средств показал их несовершенство. Генерируемые ими ответы, как правило, достаточно оперативны, но не всегда являются релевантными [Dziri et al., 2022]. В связи с этим большой интерес вызывают исследования, направленные на разработку инструментов и методов оценивания средств обработки естественного языка информации и, в частности, диалоговых систем [Lopez-Cozar et al., 2003; Wen et al., 2016; Польщиков и др., 2019; Polshchikov et al., 2019; Ji et al., 2022].

Достаточное распространение получили средства оценивания эффективности диалоговых систем, которые основаны на учете мнений людей-экспертов [Ai et al., 2008; Dodge et al., 2016; Wen et al., 2016; Польщиков и др., 2020; Polshchikov et al., 2020; Ji et al., 2022]. Использование человеческих оценок и опросников требует существенных финансовых и временных затрат, что определяет необходимость разработки новых моделей и методов, позволяющих сократить использование труда людей [Deriu et al., 2021; Dziri et al., 2022]. Для оценивания человеко-машинных диалогов применяются различные показатели, например, скорость распознавания и качество понимания системой задаваемых вопросов, естественность диалога, устойчивость системы к различным ошибкам, функциональная производительность, а также точность, лаконичность и полнота генерируемых ответов [Lopez-Cozar et al., 2003; Wen et al., 2015; Tseng et al., 2021; Xiang et al., 2021; Махди и др., 2022]. При этом вышеуказанные показатели недостаточно приспособлены к оцениванию работы вопросно-ответной системы с точки зрения удовлетворенности пользователей результатами человеко-машинного диалога. Необходимость совершенствования инструментария для оценивания характеристик диалоговых систем определяет актуальность проведенного исследования.

Цель статьи – совершенствование средств оценивания характеристик вопросно-ответной системы на основе разработки оригинальной математической модели человеко-машинного диалога.

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.

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 (что соответствует переходу в состояние «F»), то считается, что цель диалога достигнута, и моделируемый процесс переходит в терминальное состояние «R». Если в полученном ответе практически отсутствует

требуемая информация (состояние «А»), пользователь может задать системе уточняющий вопрос (состояние «Q»), а может прекратить диалог без достижения цели (терминальное состояние «N»), посчитав систему бесполезной для удовлетворения его информационной потребности. В случае получения ответа, в котором частично содержится требуемая информация, моделируемый процесс переходит в состояние «P». При этом возможны следующие варианты:

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

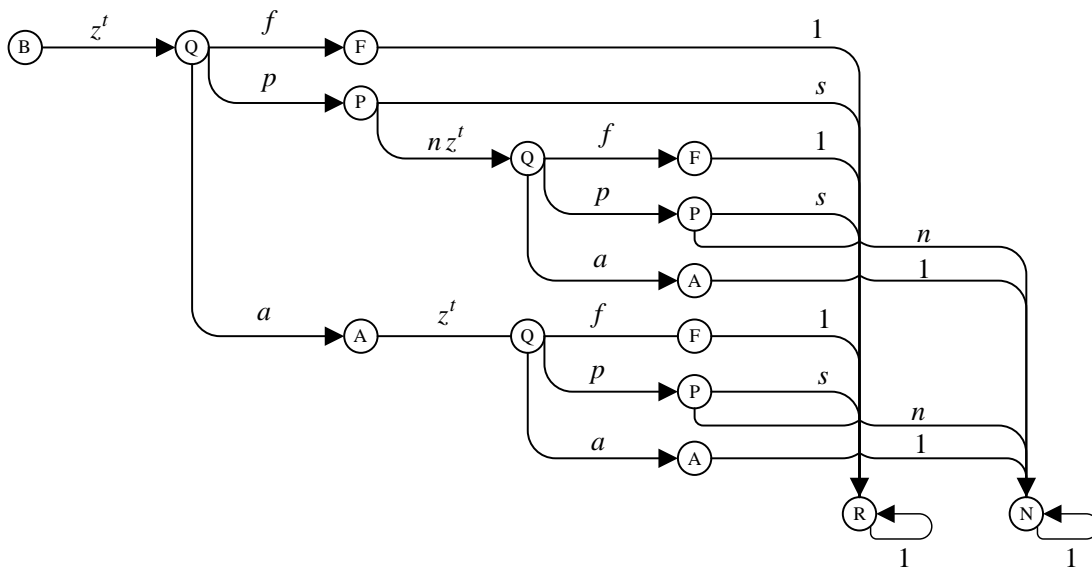


Рис. 1. Вероятностно-временной граф, моделирующий процесс человеко-машинного диалога
 Fig 1. Probabilistic-temporal graph modeling the process of human-machine dialogue

В результате выполнения конечного числа указанных выше переходов диалог завершается, то есть моделируемый процесс оказывается в терминальном состоянии «R» или в терминальном состоянии «N».

Рассматриваемый граф 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 – продолжительность по времени формулировки пользователем вопроса и выдачи системой ответа.

Переходу из состояния «Q» в состояние «F» соответствует величина f – вероятность того, что полученный пользователем ответ в полной мере содержит требуемую информацию. Переходу из состояния «Q» в состояние «P» соответствует величина p – вероятность

получения ответа, в котором частично содержится требуемая информация. Из состояния «Q» возможен также переход в состояние «A». Этому событию соответствует величина a – вероятность отсутствия в полученном ответе требуемой информации. По условию нормировки вероятность a может быть вычислена с помощью выражения:

$$a = 1 - f - p. \quad (2)$$

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 $n z^t$, где n – вероятность неудовлетворенности пользователя полученным ответом, в котором частично содержится требуемая информация. По условию нормировки вероятностей величину n можно вычислить с помощью выражения:

$$n = 1 - s. \quad (3)$$

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. Граф, изображенный на рисунке 1, может быть существенно упрощен на основе применения эквивалентных преобразований и представлен в виде, показанном на рисунке 2.

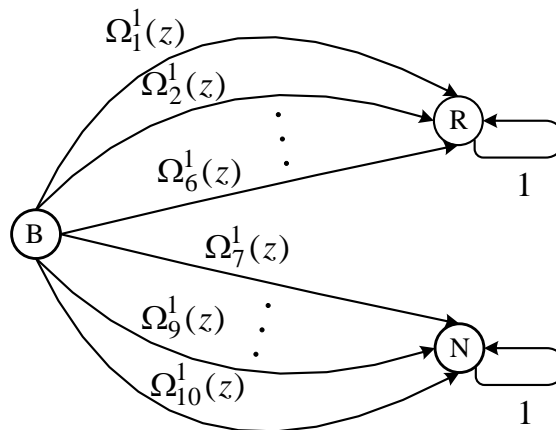


Рис. 2. Граф при $m = 1$ в преобразованном виде
 Fig 2. Graph at $m = 1$ in the transformed form

Переходы из состояний «B» в терминальное состояние «R» моделируются следующими функциями:

$$\Omega_1^1(z) = f z^t, \quad (4)$$

$$\Omega_2^1(z) = p s z^t, \quad (5)$$

$$\Omega_3^1(z) = p n f z^{2t}, \quad (6)$$

$$\Omega_4^1(z) = p n p s z^{2t}, \quad (7)$$

$$\Omega_5^1(z) = a f z^{2t}, \quad (8)$$

$$\Omega_6^1(z) = a p s z^{2t} . \quad (9)$$

Переходы из состояний «В» в терминальное состояние «N» моделируются следующими функциями:

$$\Omega_7^1(z) = p n p n z^{2t} , \quad (10)$$

$$\Omega_8^1(z) = p n a z^{2t} , \quad (11)$$

$$\Omega_9^1(z) = a p n z^{2t} , \quad (12)$$

$$\Omega_{10}^1(z) = a a z^{2t} . \quad (13)$$

Дальнейшее преобразование графа позволяет представить его в виде, изображенном на рисунке 3.

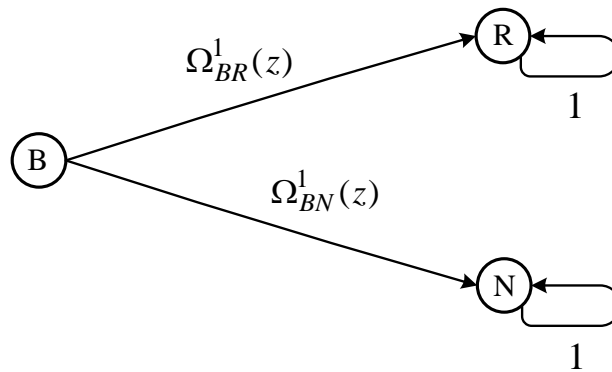


Рис. 3. Граф при $m = 1$ после эквивалентных преобразований
 Fig 3. Graph at $m = 1$ after equivalent transformations

Функция перехода из состояния «В» в терминальное состояние «R» имеет следующий вид:

$$\begin{aligned} \Omega_{BR}^1(z) &= \Omega_1^1(z) + \Omega_2^1(z) + \Omega_3^1(z) + \Omega_4^1(z) + \Omega_5^1(z) + \Omega_6^1(z) = \\ &= f z^t + p s z^t + p n f z^{2t} + p n p s^{2t} + a f z^{2t} + a p s z^{2t} . \end{aligned} \quad (14)$$

Выражение для вычисления величины $\Omega_{BR}^1(z)$ можно представить в следующем виде:

$$\begin{aligned} \Omega_{BR}^1(z) &= (f z^t + p s z^t) + (p n f z^{2t} + p n p s^{2t}) + (a f z^{2t} + a p s z^{2t}) = \\ &= (f + p s) z^t \left[(p n z^t)^0 (a z^t)^0 + (p n z^t)^1 (a z^t)^0 + (p n z^t)^0 (a z^t)^1 \right] . \end{aligned} \quad (15)$$

Функция перехода из состояния «В» в терминальное состояние «N» может быть представлена в следующем виде:

$$\begin{aligned} \Omega_{BN}^1(z) &= \Omega_7^1(z) + \Omega_8^1(z) + \Omega_9^1(z) + \Omega_{10}^1(z) = \\ &= p n p n z^{2t} + p n a z^{2t} + a p n z^{2t} + a a z^{2t} = \\ &= (p n p n z^{2t} + p n a z^{2t}) + (a p n z^{2t} + a a z^{2t}) = \\ &= (p n + a) z^{2t} \left[(p n)^0 (a)^0 + (p n)^1 (a)^0 + (p n)^0 (a)^1 \right] . \end{aligned} \quad (16)$$

После окончательных эквивалентных преобразований граф будет иметь простейший вид, представленный на рисунке 4.

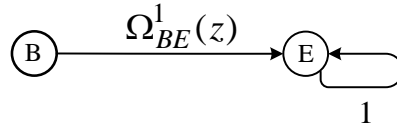


Рис. 4. Граф при $m = 1$ в простейшем виде
 Fig 4. Graph at $m = 1$ in its simplest form

На графе, изображенном на рисунке 4, состояние «Е» моделирует окончание процесса человеко-машинного диалога. Функция перехода из состояния «В» в состояние «Е» имеет следующий вид:

$$\Omega_{BE}^1(z) = \Omega_{BR}^1(z) + \Omega_{BN}^1(z) \quad (17)$$

Чтобы вывести универсальные формулы для вычисления вероятности достижения цели человеко-машинного диалога и его средней продолжительности, нужно получить выражения для вычисления функций $\Omega_{BR}^m(z)$, $\Omega_{BN}^m(z)$ и $\Omega_{BE}^m(z)$ при любом натуральном m . С этой целью были выведены эти выражения для $m = 2$, $m = 3$ и $m = 4$, а затем в них выявлены обобщающие закономерности.

Рассмотрим выражение (15). Значение степени, в которую возводятся множители $(pn z^t)$, обозначим α . Значение степени, в которую возводятся множители $(a z^t)$, обозначим β . Число слагаемых, в которых множитель $(pn z^t)$ возведен в степень α и множитель $(a z^t)$ возведен в степень β , обозначим γ . В результате разработки математических моделей процесса человеко-машинного диалога при $m = 1$, $m = 2$, $m = 3$ и $m = 4$ получены данные о значениях величины γ в зависимости от того, какие значения принимают величины α и β . Анализ этих данных позволил выявить обобщающие закономерности в моделировании исследуемого процесса и представить универсальное выражение для вычисления величины $\Omega_{BR}^m(z)$ при любом натуральном m в следующем виде:

$$\Omega_{BR}^m(z) = (f + ps)z^t \sum_{\alpha=0}^m \sum_{\beta=0}^{m-\alpha} (pnz^t)^\alpha (\alpha z^t)^\beta \frac{(\alpha + \beta)!}{\alpha! \beta!}. \quad (18)$$

Рассмотрим выражение (16). Значение степени, в которую возводятся множители (pn) , обозначим η . Тогда значение степени, в которую возводятся множители a , равны $(m - \eta)$. Число слагаемых, в которых множитель (pn) возведен в степень η и множитель a возведен в степень $(m - \eta)$, обозначим μ . В результате разработки математических моделей процесса человеко-машинного диалога при $m = 1$, $m = 2$, $m = 3$ и $m = 4$ получены данные о значениях величины μ в зависимости от того, какие значения принимают величины η и $(m - \eta)$. Анализ этих данных позволил выявить обобщающие закономерности в моделировании исследуемого процесса и представить универсальное выражение для вычисления величины $\Omega_{BN}^m(z)$ при любом натуральном m в следующем виде:

$$\Omega_{BN}^m(z) = (pn + a)z^{(m+1)t} \sum_{\alpha=0}^m (pn)^\alpha a^{(m-\alpha)} \frac{m!}{\alpha!(m-\alpha)!}. \quad (19)$$

Функция перехода из состояния «В» в состояние «Е» при любом натуральном m вычисляется с помощью следующего выражения:

$$\Omega_{BE}^m(z) = \Omega_{BR}^m(z) + \Omega_{BN}^m(z) \quad (20)$$

Тогда для вычисления вероятности достижения цели диалога можно использовать выражение:

$$r_m = \Omega_{BR}^m(z) \Big|_{z=1} \quad (21)$$

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

$$u_m = \frac{d}{dz} \Omega_{BE}^m(z) \Big|_{z=1} \quad (22)$$

Проведение вычислительных экспериментов

На основе применения представленных выше выражений проведены вычислительные эксперименты по оцениванию вероятностно-временных характеристик человеко-машинного диалога. На рисунках 5 и 6 представлены полученные графики зависимости величин r_m и u_m соответственно от значений величины m при $f = 0,3$ и $p = 0,4$. Анализ графиков показывает, что по мере увеличения допустимого числа задаваемых системе уточняющих вопросов вероятность достижения цели человеко-машинного диалога повышается и приближается к 1, однако при этом существенно возрастает средняя продолжительность диалога.

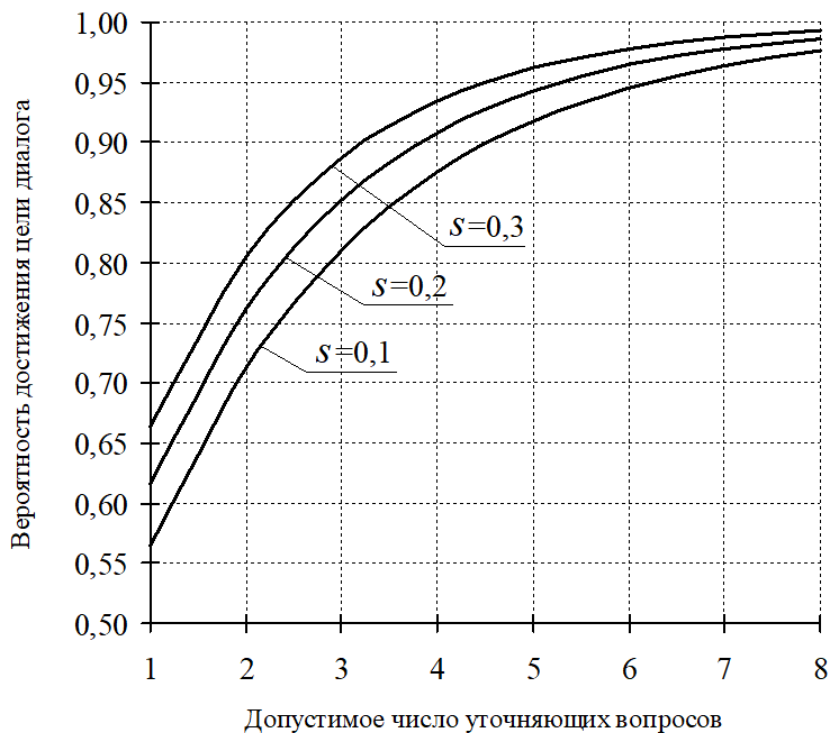


Рис. 5. Графики зависимости величины r_m от значений величины m при $f = 0,3$ и $p = 0,4$

Fig 5. Graphs of the dependence of the quantity r_m on the values m at $f = 0,3$ and $p = 0,4$

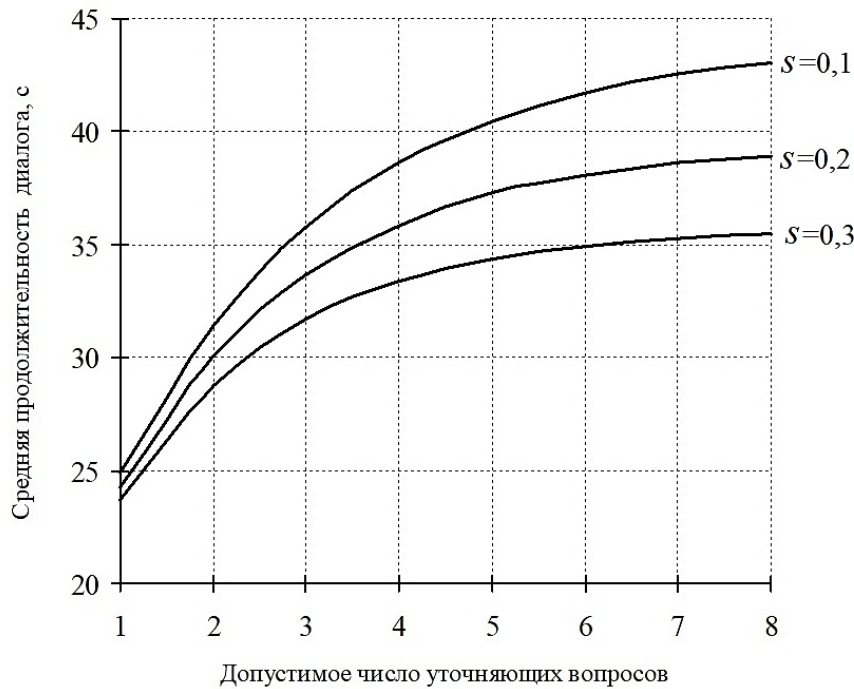


Рис. 6. Графики зависимости величины u_m от значений величины m при $f=0,3$ и $p=0,4$

Fig 6. Graphs of the dependence of the quantity u_m on the values m at $f=0,3$ and $p=0,4$

Пусть 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 5 и 6 графиков, показывает, что диалоговая система с параметрами $f=0,3$ и $p=0,4$ рекомендуется тем пользователям, которые в случае необходимости готовы задать 5 уточняющих вопросов, а также с вероятностью не ниже 0,2 будут удовлетворены полученными ответами, в которых лишь частично содержится требуемая информация.

Таким образом, разработанная математическая модель позволяет обосновать целесообразность применения вопросно-ответной системы с конкретными характеристиками f и p для удовлетворения информационных потребностей тех пользователей, которым свойственны определенные значения m и s .

Заключение

Представленные в статье результаты исследования позволяют усовершенствовать средства оценивания характеристик осуществляемого на естественном языке человеко-машинного диалога. Разработана модель процесса человеко-машинного диалога на основе применения математического аппарата вероятностно-временных графов. Модель ориентирована на оценивание диалога с точки зрения достижения его цели, состоящей в удовлетворении определенной информационной потребности пользователя. С помощью модели можно вычислить вероятность достижения цели диалога и его среднюю продолжительность с учетом характеристик вопросно-ответной системы и значений показателей, характеризующих особенности пользователя. В качестве характеристик вопросно-ответной системы используются: 1) вероятность того, что выданный системой ответ в полной мере содержит информацию, которая требуется пользователю; 2) вероятность выдачи системой ответа, в котором лишь частично содержится требуемая пользователю информация. Особенности пользователя характеризуются следующими величинами: 1) предельное число уточняющих вопросов, которое, в случае необходимости,

пользователь готов задать системе; 2) вероятность удовлетворенности пользователя полученным ответом, в котором частично содержится требуемая информация. Применение модели позволяет обосновать выбор диалоговых систем с конкретными характеристиками и рекомендовать их тем или иным группам пользователей для удовлетворения информационных потребностей.

Дальнейшие исследования в рамках представленной тематики планируется посвятить разработке методики оценивания характеристик вопросно-ответных систем и значений показателей, характеризующих особенности пользователей этих систем.

Список литературы

- Агузумян Р.В., Великанова А.С., Польщиков К.А., Игитян Е.В., Лихошерстов Р.В. 2021. О применении интеллектуальных технологий обработки естественного языка и средств виртуальной реальности для поддержки принятия решений при подборе исполнителей проектов. *Экономика. Информатика*, 48(2): 392–404. DOI 10.52575/2687-0932-2021-48-2-392-404.
- Махди Т.Н., Игитян Е.В., Польщиков К.А., Корсунов Н.И. 2022. Оценивание эффективности функционирования диалоговой системы на основе применения нечеткого вывода с нейросетевой настройкой. *Экономика. Информатика*, 49(2): 356–374. DOI 10.52575/2687-0932-2022-49-2-356-374.
- Польщиков К.А., Лазарев С.А., Константинов И.С., Польщикова О.Н., Свойкина Л.Ф., Игитян Е.В., Балакшин М.С. 2020. Модель для оценки эффективности выполнения робототехнической системой коммуникативных функций. *СТИН*, 6: 4–7.
- Польщиков К.А., Польщикова О.Н., Игитян Е.В., Балакшин М.С. 2019. Алгоритм поддержки принятия решений по выбору средств обработки больших массивов естественно-языковых данных. *Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика*. 46(3): 553–562. DOI 10.18413/2411-3808-2019-46-3-553-562.
- Ai H., Litman D.J. 2008. Assessing Dialog System User Simulation Evaluation Measures Using Human Judges. *Proceedings of ACL-08*: 622–629.
- Deriu J., Rodrigo A., Otegi A. 2021. Survey on evaluation methods for dialogue systems. *Artificial Intelligence Review*, 54: 755–810.
- Dodge J., Gane A., Zhang X., Bordes A., Chopra S., Miller A., Szlam A., Weston J. 2016. Evaluating Prerequisite Qualities for Learning End-to-End Dialog Systems. *arXiv*. URL: <https://doi.org/10.48550/arXiv.1511.06931> (accessed: 15.02.2023).
- Dziri N., Rashkin H., Linzen T., Reitter D. 2022. Evaluating Attribution in Dialogue Systems: The BEGIN Benchmark. *Transactions of the Association for Computational Linguistics*, 10: 1066–1083.
- Ji T., Graham Y., Jones G.J.F., Lyu C., Liu Q. 2022. Achieving Reliable Human Assessment of Open-Domain Dialogue Systems. *arXiv*. URL: <https://doi.org/10.48550/arXiv.2203.05899> (accessed: 15.02.2023).
- Lopez-Cozar R., De la Torre A., Segura J.C., Rubio A.J. 2003. Assessment of dialogue systems by means of a new simulation technique. *Speech Communication*, 40(3): 387–407.
- Polshchikov K., Lazarev S., Polshchikova O., Igityan E. 2019. The Algorithm for Decision-Making Supporting on the Selection of Processing Means for Big Arrays of Natural Language Data. *Lobachevskii Journal of Mathematics*, 40(11): 1831–1836.
- Polshchikov K.A., Lazarev S.A., Konstantinov I.S., Polshchikova O.N., Svoikina L.F., Igityan E.V., Balakshin M.S. 2020. Assessing the Efficiency of Robot Communication. *Russian Engineering Research*, 40: 936–938.
- Polshchikov K.A., Velikanova A.S., Igityan E.V. 2022. Neural network natural language processing tools for identifying personal priorities in the project performers selection in the field of smart agriculture. *IOP Conference Series: Earth and Environmental Science*, 1069: 012012.
- Shah H., Warwick K., Vallverdú J., Wu D. 2016. Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, 58: 278–295.
- Tseng B.-H., Dai Y., Kreyszig F., Byrne B. 2021. Transferable Dialogue Systems and User Simulators. *arXiv*. URL: <https://doi.org/10.48550/arXiv.2107.11904> (accessed: 15.02.2023).
- Velikanova A.S., Polshchikov K.A., Likhoshertov R.V., Polshchikova A.K. 2021. The use of virtual reality and fuzzy neural network tools to identify the focus on achieving project results. *Journal of*

Physics: Conference Series. 2nd International Scientific Conference on Artificial Intelligence and Digital Technologies in Technical Systems 2021, Volgograd, 2060: 173707.

- Wen T.-H., Gasic M., Mrksic N., Rojas-Barahona L.M., Su P.-H., Vandyke D., Young S. 2016. Multi-domain Neural Network Language Generation for Spoken Dialogue Systems. arXiv. URL: <https://doi.org/10.48550/arXiv.1603.01232> (accessed: 15.02.2023).
- Wen T.-H., Gasic M., Mrksic N., Su P.-H., Vandyke D., Young S. 2015. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. arXiv. URL: <https://doi.org/10.48550/arXiv.1508.01745> (accessed: 15.02.2023).
- Wen T.-H., Vandyke D., Mrksic N., Gasic M., Rojas-Barahona L.M., Su P.-H., Ultes S., Young S. 2017. A Network-based End-to-End Trainable Task-oriented Dialogue System. arXiv. URL: <https://doi.org/10.48550/arXiv.1604.04562> (accessed: 15.02.2023).
- Xiang J., Liu Y., Cai D., Li H., Lian D., Liu L. 2021. Assessing Dialogue Systems with Distribution Distances. arXiv. URL: <https://doi.org/10.48550/arXiv.2105.02573> (accessed: 15.02.2023).

References

- Aguzumtsyan R.V., Velikanova A.S., Polshchikov K.A., Igityan E.V., Likhosherstov R.V. 2021. Application of intellectual technologies of natural language processing and virtual reality means to support decision-making when selecting project executors. *Economics. Information technologies*, 48(2): 392–404. (in Russian). DOI 10.52575/2687-0932-2021-48-2-392-404.
- Mahdi T.N., Igityan E.V., Polshchikov K.A., Korsunov N.I. 2022. Evaluation of the Dialogue System Efficiency Based on the Application of Fuzzy Inference with Neural Network Settings. *Economics. Information technologies*, 49(2): 356–374 (in Russian). DOI 10.52575/2687-0932-2022-49-2-356-374/
- Polshchikov K.A., Lazarev S.A., Konstantinov I.S., Polshchikova O.N., Svoikina L.F., Igityan E.V., Balakshin M.S. 2020. A model for evaluating the effectiveness of a robotic system performing communicative functions. *STIN*, 6: 4-7 (in Russian).
- Polshchikov K.A., Polshchikova O.N., Igityan E.V., Balakshin M.S. 2019. The algorithm of decision support in the choice of means of processing large amounts of natural language data. *Belgorod State University Scientific Bulletin. Economics. Information technologies*. 46(3): 553–562 (in Russian). DOI 10.18413/2411-3808-2019-46-3-553-562.
- Ai H., Litman D.J. 2008. Assessing Dialog System User Simulation Evaluation Measures Using Human Judges. *Proceedings of ACL-08*: 622–629.
- Deriu J., Rodrigo A., Otegi A. 2021. Survey on evaluation methods for dialogue systems. *Artificial Intelligence Review*, 54: 755–810.
- Dodge J., Gane A., Zhang X., Bordes A., Chopra S., Miller A., Szlam A., Weston J. 2016. Evaluating Prerequisite Qualities for Learning End-to-End Dialog Systems. arXiv. URL: <https://doi.org/10.48550/arXiv.1511.06931> (accessed: 15.02.2023).
- Dziri N., Rashkin H., Linzen T., Reitter D. 2022. Evaluating Attribution in Dialogue Systems: The BEGIN Benchmark. *Transactions of the Association for Computational Linguistics*, 10: 1066–1083.
- Ji T., Graham Y., Jones G.J.F., Lyu C., Liu Q. 2022. Achieving Reliable Human Assessment of Open-Domain Dialogue Systems. arXiv. URL: <https://doi.org/10.48550/arXiv.2203.05899> (accessed: 15.02.2023).
- Lopez-Cozar R., De la Torre A., Segura J.C., Rubio A.J. 2003. Assessment of dialogue systems by means of a new simulation technique. *Speech Communication*, 40(3): 387–407.
- Polshchikov K., Lazarev S., Polshchikova O., Igityan E. 2019. The Algorithm for Decision-Making Supporting on the Selection of Processing Means for Big Arrays of Natural Language Data. *Lobachevskii Journal of Mathematics*, 40(11): 1831–1836.
- Polshchikov K.A., Lazarev S.A., Konstantinov I.S., Polshchikova O.N., Svoikina L.F., Igityan E.V., Balakshin M.S. 2020. Assessing the Efficiency of Robot Communication. *Russian Engineering Research*, 40: 936–938.
- Polshchikov K.A., Velikanova A.S., Igityan E.V. 2022. Neural network natural language processing tools for identifying personal priorities in the project performers selection in the field of smart agriculture. *IOP Conference Series: Earth and Environmental Science*, 1069: 012012.
- Shah H., Warwick K., Vallverdú J., Wu D. 2016. Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, 58: 278–295.

- Tseng B.-H., Dai Y., Kreyszig F., Byrne B. 2021. Transferable Dialogue Systems and User Simulators. arXiv. URL: <https://doi.org/10.48550/arXiv.2107.11904> (accessed: 15.02.2023).
- Velikanova A.S., Polshchikov K.A., Likhoshevstov R.V., Polshchikova A.K. 2021. The use of virtual reality and fuzzy neural network tools to identify the focus on achieving project results. Journal of Physics: Conference Series. 2nd International Scientific Conference on Artificial Intelligence and Digital Technologies in Technical Systems 2021, Volgograd, 2060: 173707.
- Wen T.-H., Gasic M., Mrksic N., Rojas-Barahona L.M., Su P.-H., Vandyke D., Young S. 2016. Multi-domain Neural Network Language Generation for Spoken Dialogue Systems. arXiv. URL: <https://doi.org/10.48550/arXiv.1603.01232> (accessed: 15.02.2023).
- Wen T.-H., Gasic M., Mrksic N., Su P.-H., Vandyke D., Young S. 2015. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. arXiv. URL: <https://doi.org/10.48550/arXiv.1508.01745> (accessed: 15.02.2023).
- Wen T.-H., Vandyke D., Mrksic N., Gasic M., Rojas-Barahona L.M., Su P.-H., Ultes S., Young S. 2017. A Network-based End-to-End Trainable Task-oriented Dialogue System. arXiv. URL: <https://doi.org/10.48550/arXiv.1604.04562> (accessed: 15.02.2023).
- Xiang J., Liu Y., Cai D., Li H., Lian D., Liu L. 2021. Assessing Dialogue Systems with Distribution Distances. arXiv. URL: <https://doi.org/10.48550/arXiv.2105.02573> (accessed: 15.02.2023).

Конфликт интересов: о потенциальном конфликте интересов не сообщалось.

Conflict of interest: no potential conflict of interest related to this article was reported.

ИНФОРМАЦИЯ ОБ АВТОРАХ

INFORMATION ABOUT THE AUTHORS

Игитян Елена Владимировна, аспирант кафедры информационно-телекоммуникационных систем и технологий Белгородского государственного национального исследовательского университета, г. Белгород, Россия

Elena V. Igityan, Post-graduate Student of the Department of Information and Telecommunications Systems and Technologies of the Belgorod National Research University, Belgorod, Russia

Польщикова Константин Александрович, доктор технических наук, доцент, директор института инженерных и цифровых технологий Белгородского государственного национального исследовательского университета, г. Белгород, Россия

Konstantin A. Polshchikov, Doctor of Technical Sciences, Associate Professor, Director of the Institute of Engineering and Digital Technologies of the Belgorod National Research University, Belgorod, Russia

Немцев Александр Николаевич, кандидат физико-математических наук, доцент, доцент кафедры прикладной информатики и информационных технологий Белгородского государственного национального исследовательского университета, г. Белгород, Россия

Alexander N. Nemtsev, Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Applied Informatics and Information Technologies, Belgorod National Research University, Belgorod, Russia