

**Х.Б. Штанчаев****СТАТИСТИЧЕСКИЕ И МАШИННЫЕ МЕТОДЫ АВТОМАТИЧЕСКОГО  
ИЗВЛЕЧЕНИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ ИЗ ТЕКСТА  
(ОБЗОР)**

*Для решения задачи автоматического извлечения причинно-следственных связей (ПСС) до 2000-х годов использовалась концепция нестатистических методов. Данные методы использовали построенные вручную лингвистические шаблоны. Очевидно, что ПСС, которые не подходили в построенные шаблоны, могли быть не определены. Нестатистические методы требовали постоянного ручного контроля со стороны экспертов, вплоть до оценки. Практически все методы были направлены на извлечение явных ПСС. В некоторых методах были использованы попытки отвязать систему извлечения от конкретной предметной области. Для исключения вышеупомянутых недостатков, разрабатываемые в дальнейшем методы начали смещаться в сторону обработки статистических данных и машинного обучения. Был проанализирован целый ряд ценных работ, связанных с новой парадигмой извлечения ПСС. Целью исследования было оценить новые методы с возможностью выявить их преимущества и недостатки. Большим преимуществом машинных и статистических методов является независимость от предметной области с сохранением точности извлечения. Такие методы хуже по точности, однако не привязаны к конкретной проблемной области. Сами методы в отличие от нестатистических, которые использовали лингвистическое и синтаксическое сравнение с шаблонами вручную, сосредоточены на поиске этих самых шаблонов. Несмотря на то, что машинные и статистические методы в своем большинстве являются независимыми от предметной области и для обучения используют большие корпуса текста они предназначены преимущественно для английского языка. Так же отсутствует стандартизированный набор данных, который позволил бы сравнить методы между собой. Все работы, посвященные методам, проигнорировали извлечение неявных ПСС.*

*Причинно-следственные связи; причинные знания; обработка естественного языка; машинное обучение; компьютерная лингвистика; скрытые причинные связи.*

**Kh.B. Shtanchaev****STATISTICAL AND MACHINE METHODS FOR AUTOMATICALLY  
EXTRACTING CAUSAL RELATIONSHIPS FROM TEXT (REVIEW)**

*Until the 2000s, the concept of non-statistical methods was used to solve the problem of automatic extraction of causal relationships (CR). These methods used manually constructed linguistic templates. Obviously, the CR that did not fit into the built templates could not be defined. Non-statistical methods required constant manual control by experts, up to the evaluation. Almost all methods were aimed at extracting explicit CR. In some methods, attempts were made to untie the extraction system from a specific subject area. To eliminate the above disadvantages, the methods developed in the future began to shift towards statistical data processing and machine learning. In this article, statistical and machine methods of CR extraction are considered. A few valuable papers related to the new paradigm of CR extraction were analyzed. The aim of the research was to evaluate new methods with the ability to identify their advantages and disadvantages. The great advantage of machine and statistical methods is independence from the subject area while maintaining the accuracy of extraction. Such methods are worse in accuracy, but they are not tied to a specific problem area. The methods themselves, unlike non-statistical ones, which used linguistic and syntactic comparison with templates manually, are focused on finding these templates. Even though machine and statistical methods are mostly independent of the subject area and use large corpora of*

*text for teaching, they are intended mainly for the English language. There is also no standardized data set that would allow methods to be compared with each other. All works devoted to methods ignored the extraction of implicit CR.*

*Causality; causal knowledge; natural language processing; machine learning; computational linguistics; hidden causality.*

**Введение.** Задача автоматического извлечения данных из слабоструктурированных текстов является разновидностью поиска и одной из проблем обработки естественного языка в настоящее время. С каждым днем объемы неструктурированной информации увеличивается стремительно. Практически во всех областях деятельности люди сталкиваются с вопросами о причинно-следственных связях, то есть связях между событиями или фактами, в которых одно является причиной, а другое – следствием. Извлечение причинно-следственных связей из текста имеет важное значение во многих областях, таких как анализ данных, экономика, медицина и машинное обучение. Из-за нечеткого и многообразного характера языка задача извлечения причинно-следственных связей из текста может быть достаточно сложной. Одним из стандартных подходов к извлечению причинно-следственных связей является использование нестатистических методов. Обзор и анализ нестатистических методов приведен в работе [1]. Методы показали свою эффективность при извлечении ПСС. Однако осталась привязка к конкретным предметным областям и языку. Другой подход заключается в попытках использования машинного обучения.

**Статистические и машинные методы извлечения ПСС.** Ранние попытки внедрить машинное обучение для решения достаточно простых задач извлечения ПСС была предпринята Гирджу [2]. Она стала достаточно примитивной модификацией работы Гирджу и Молдована [3]. Данная работа подробно рассмотрена в обзоре нестатистических методов [1]. В своей попытке автор попытался автоматизировать ручную процедуру проверки и ранжирования шаблонов для каждого составленного шаблона, использовав метод деревьев решений C4.5, разработанный Джоном Квинланом [4]. Алгоритм, состоящий из двух этапов, использовался для определения достаточно простых семантических сетей слов:

$$\langle \text{NP}_1 \text{ verb NP}_2 \rangle, \quad (1)$$

где verb – причинный глагол, NP – существительное или фраза существительное плюс другая часть речи.

Первый этап алгоритма – определение лексико-синтаксической закономерности автор разделил на три шага:

1. Подбор семантического отношения для работы модели. Автор указал, что в рамках своей работы семантическое правило будет определено как «причинность».

2. Определение языковой конструкции  $C_i$  и  $C_j$ , которые содержат семантическое правило. Для определения языковых конструкций использовалась лексическая база данных Wordnet.

3. Извлечение лексико-семантической взаимосвязи, которая будет содержать  $C_i$  и  $C_j$ .

Второй этап алгоритма – обучение модели машинного обучения. Для обучения использовался языковой корпус TREC 9. Для тренировочного набора использовался следующий вид записи:

$$\langle \text{causeNP}; \text{verb}; \text{effectNP}; \text{target} \rangle,$$

где *causeNP* – причина, *verb* – глагол, *effectNP* – следствие, *target* – может быть «Да» или «Нет» в зависимости является ли утверждение истинным или ложным. Используя данный тренировочный набор, разделенный на 10 блоков перекрестной проверки (кросс-валидации) методом дерева решений, авторы обучили модель. Точность и полнота работы модели составили 73,91% и 88,69% на учебном наборе из 6000 предложений и тестовом наборе из 1200 предложений. Для наборов не была определена конкретная предметная область. Впервые в качестве семантических признаков автором была предложена идея использования семантической сети слов. Данные указывают на то, что работа автора с использованием машинного обучения оказалась гораздо эффективной при извлечении ПСС чем оригинальная работа [3]. Можно предположить, что ошибки полученные в процессе обучения и проверки системы связаны с малым количеством причинных глаголов (около 60) и с недостаточным набором данных. Решение задачи с малым количеством данных, с которой встретился автор [2] можно было бы решить, увеличив наборы данных, однако это требует очень много ресурсов.

Очень ценная работа с точки зрения подхода к классификации семантических отношений получилась у Марку и Эчихаби [5]. Для определения различных отношений использовалась Байесовская классификация [6]. Модель не стала самостоятельной, но в дальнейшем работа была модифицирована уже другими авторами. Примером такой модификации может служить работа Чанг и Чой [7]. Авторы добавили анализатор зависимостей [8], который извлекает выражения вида «*NP1 causal NP2*». Вследствие чего из текста извлекались вероятности этих троек. Найденные выражения фильтровались с помощью набора причинных глаголов (60 глаголов определенных еще Гирджу [2]) по убывающей вероятности.

«Кандидат в причинность» [7] вышеописанная тройка  $t_i$ , классифицировалась как причинная  $c_1$  или не причинная  $c_0$ . Для решения задачи классификации авторы использовали классификатор основанный на байесовском:

$$c^* = \arg \max_{c_j} P(c_j | t_i) = \operatorname{argmax}_{c_j} \frac{P(c_j)P(t_i | c_j)}{P(t_i)}. \quad (2)$$

Он классифицировал тройку с большими вероятностями как «причинную» или как «не причинную». Для улучшения вероятности используется процедура максимизации математического ожидания. Для оценки работы метода авторы использовали не зависящий от предметной области текст в 5 млн предложений. Обучающий набор так же предполагал вручную определенные два текста. Авторы не указывают размер текстов, но точно сообщают, что один текст был независимый от предметной области, а другой связанный с медициной. Метод показал точность 82,88%, полноту – 64,79%, а F-score 72.73%. Так же в статье авторы продемонстрировали как метод еще может быть улучшен. Вся работа была направлена на улучшение точности, в том числе и применение байесовских классификаторов. В работе было продемонстрировано как извлекать ПСС на основе статистических моделей на больших данных.

Бланко в 2008 году предложил работу [9] для извлечения явных ПСС. Для извлечения автор использовал синтаксические шаблоны. Авторы вручную классифицировали 1270 предложений из корпуса TREC5 определяющую или не определяющую причинно-следственную связь. Было обнаружено 170 промежуточных причинно-следственных связей. Предложения, определяющие причинно-следственную связь, были вручную сгруппированы по синтаксическим шаблонам показанным в табл. 1.

Таблица 1

Синтаксические шаблоны, выражающие причинно-следственную связь, их продуктивность и примеры

№	Шаблон	Эффект., %	Пример (англ. язык)
1	[VP rel C], [rel c, VP]	63,75	We didn't go because it was raining
2	[NP VP NP]	13,75	The speech sparked a controversy
3	[VP rel NP], [rel NP, VP]	8.12	He died of cancer
4	другие.	14.38	The lightning caused the workers to fall

Самый популярный шаблон представлял собой следующую последовательность:

$$[VP \text{ rel } C], [rel, VP], \quad (3)$$

где VP – глагольная фраза, rel – соединитель (связующий), C – причинный глагол.

«Соединитель(связующий), который может быть либо предлогом, либо союзом. Ручная кластеризация позволила понять, что четырьмя наиболее распространенными соединителями, определяющими причинно-следственную связь, являются *after, as, be cause* и *since*» [9].

Для обучения авторы использовали языковой корпус. 1068 экземпляров фраз были разделены на обучающую (75%) и тестовую (25%) выборки. В качестве алгоритма обучения использовалась реализация пакетирования с деревьями решений C4.5. Авторы получили средний показатель F-score, равный 0,91, по сравнению с 0,7 у Гирджу [2] и 0,72 у Чанга [7].

Позже, в 2010 году Сил и другие в своей работе [10] реализовали извлечение ПСС с их предпосылками и с постусловиями. Предполагалось, что такой подход будет предполагать извлечение ПСС «с точки зрения извлечения знаний здравого смысла из текста». В системе, которая была названа PREPOST, любое действие осуществляется сбор большого набора содержащих слова согласно шаблону «*is/are/was/were A-ing*». Где A – слово в тексте D<sub>i</sub>. Из этого большого набора создается список возможных предусловий и постусловий путем вычисления поточечной вероятности (PMI) [11]:

$$PMI(A, w) = \log \frac{|d \in D_A | w \text{ appears in } d|}{|D_A| * |d \in D | w \text{ appears in } d|} \quad (4)$$

Как показали тесты система PREPOST приспособлена только для небольших наборов данных только для типа ПСС, попадающих в группу «событие – состояние». Однако система хорошо определяет скрытые ПСС с высокой точностью. Точность метода составила 94%.

Вышеописанные методы не решили проблему неявных ПСС. Они в большей степени были сосредоточены около неоднозначности и пытались решить данную проблему во время обучения, подбирая различные функции алгоритма обучения.

Подход к проблеме поменяли Бетард и Мартин [13]. Авторы построили систему, которая «отвечала на вопрос: учитывая два события происходящих в одном предложении, можно ли рассматривать одно событие как причину другого?» [13]. В своей системе авторы использовали метки предложений:

<sup>1</sup> В англоязычной литературе «event- state». Например: «увеличение давления привело к аварии».

**ДО (before)** – предложения, в которых событие полностью предшествует второму;

**ПОСЛЕ (after)** – предложения, в которых второе событие полностью предшествует первому;

**БЕЗ-СОБЫТИЯ(No-Rel)** – ни одно из событий явно не предшествует другому.

Так же авторы решили другую задачу, не решавшуюся в предыдущих работах. Это разделение связей на временные и причинные. Используя корпус от Google, «поверхностные функции» [13], они получили оценку F-score равную 0,49 для временных связей и 0,524 для ПСС.

**Графовые методы извлечения ПСС.** При решении задач извлечения ПСС не обошлось и без графовых моделей и методов. Ринк в своей работе [14] предложил графовую модель для извлечения ПСС. На основании текстовых предложений строился граф. В качестве вершин графа выступали лексемы. Каждый такой граф мог содержать информацию о лексической и синтаксической структуре предложения. ПСС извлекались из графа в виде подграфов определенным алгоритмом [15]. Для повышения точности используются различные цепочки и семантические связи между глаголами. Используя языковой корпус от Google, что и предыдущий автор, метод показал F-score равной 0,579. Что лучше, чем работа Бетарда и Мартина [13]. Это означает, что графовый метод имеет преимущество, которое заключается в постоянном удерживании информации всего предложения как в части лексики, так и синтаксиса. Кроме того, удерживается весь контекст и взаимосвязь между всеми объектами. Остальная доля ошибок вновь была связана с двусмысленностями.

Достаточно сложные методы решения поставленной задачи предложили Соргенте [16], а также Ян и Мао [17].

В своей работе Соргенте[16] из предложения S извлекает наборы пар причинных связей:

$$\{(C_1E_1), (C_2E_2), \dots, (C_nE_n)\}, \quad (5)$$

где  $(C_iE_i)$  – i-ю причинную связь в предложении S.

Для этого используется метод, предложенный автором, алгоритм которого можно увидеть на рис. 1.

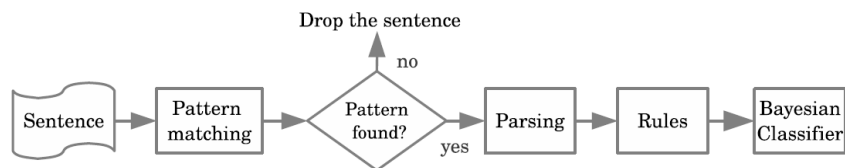


Рис. 1. Алгоритм извлечения причинных пар, предложенный автором [16]

В первую очередь предложение проверяется на содержание одного из шаблонов. Для поиска причинных пар в предложении используются шаблоны, основываясь на языковых конструкциях представленных в табл. 2.

Таблица 2

## Соответствие языковой конструкции и шаблона для поиска причинной связи в предложении

Языковая конструкция	Шаблон для поиска причинной связи (английский язык)	Шаблон для поиска причинной связи (перевод на русский язык)
Простые причинные глаголы	<cause generate triggers ...>	<причина порождение побуждение ...>
Пассивные причинные глаголы	<caused generated triggered ...> by	<явилось причиной порождено побудило> <чего-либо или чем-либо>
Комбинация «существительное-предлог»	<result cause lead ...> <in of to>	<результат причина приводить ...> <в из к>
Одиночные предлоги	<from after ...>	<из после ..>

И затем, если шаблон найден, к данному предложению применяются правила. В зависимости от того какую языковую конструкцию автор пытался определить он применял соответствующее правило. Главное правило, которое применяется для определения причинно-следственной связи, выглядит следующим образом:

$$cause(S, P, C)effect(S, P, E) \rightarrow cRel(S, C, E), \quad (6)$$

где  $cause(S, P, C)$  – обозначения того, что  $C$  является причиной в предложении  $S$  в соответствии с шаблоном  $P$ ,  $effect(S, P, E)$  – обозначение того, что  $E$  это результат в  $C$  относительно шаблона  $P$ .

Основываясь на опыте предыдущих ученых, автор использовал в своей работе байесовский классификатор для фильтрации результатов. Проводимые эксперименты на данных SemEval 2010 автор получил результат F-score равный 0,64.

Касательно работы Яна и Мао [17] систему которых назвали MLRE, можно отметить, что система представляет собой многоуровневый анализатор отношений в предложении. Система позволяет обнаруживать ПСС на основе лингвистических знаний грамматики. Следует отметить, что для обучения и тестирования используется лексико-семантический ресурс WordNet. Несмотря на то, что наборы данных приходится помечать вручную и снабжать комментариями метод показал лучший результат по метрике F-score равной 0,66.

После вышеописанных были еще работы [18–21], однако они не участвовали в улучшении статистического или машинного подхода. Скорее они использовали машинные модели извлечения в прикладных задачах, таких как извлечение ПСС между лекарственными препаратами и вирусами или извлечение ПСС из поисковых запросов для улучшения релевантности выдаваемых ответов поисковыми системами.

**Сравнение результатов работы различных методов и подходов.** Большинство авторов статистических и машинных методов извлечения ПСС использовали три основные метрики для оценки результатов работы своих методов и моделей:

1) Точность (Precision)

$$Precision = \frac{TP}{(TP+FP)}; \quad (7)$$

2) Полнота (Recall)

$$Recall = \frac{TP}{(TP+FN)}; \quad (8)$$

3) Среднее гармоническое от точности и полноты (F-score)

$$F - score = \frac{2*TP}{(2*TP+FN+FP)}. \quad (9)$$

Также следует отметить, что авторы использовали различные корпуса текстов и датасетов. Учитывая все это можно построить сводную таблицу:

Таблица 3

**Сводная таблица по статистическим и машинным методам извлечения ПСС**

№	Автор работы	Набор данных	Precision, %	Recall, %	F-score, %
1	Гирджу [2]	TREC9	73,91	<b>88,69</b>	70,6
2	Чанг и Чой [7]	TREC	82.88	64,79	72,73
3	Бланко [9]	TREC5	-	-	<b>91.3</b>
4	Сил [10]	-	<b>94,3</b>	85,4	-
5	Бетард [13]	Google N-gram	-	-	49
6	Ринк [14]	Google N-gram	-	-	57
7	Соргенте [16]	SemEval	71	58	63
8	Мао [17]	WordNet	-	-	66

**Выводы.** В статье представлен обзор исследовательской литературы, по статистическим и машинным методам автоматического извлечения причинно-следственных. Также в статье приведен подробный анализ высоко цитируемых исследований по данной тематике.

Анализ позволил сделать следующие выводы:

1. В большинстве статистические и машинные методы используют байесовский классификатор, который позволяет рассчитывать поточечную информации (PMI). Такая практика позволяет избавить методы от каких-либо ранжирующих факторов как в нестатистических методах, где используются ручные или полуавтоматические ранжирующие меры.
2. В отличие от нестатистических методов, которые использовали лингвистическое и синтаксическое сравнение с шаблонами которые вручную разрабатывались, методы машинного извлечения ПСС сосредоточены на поиске этих самых шаблонов.
3. Большим преимуществом машинных и статистических методов является независимость от предметной области с сохранением точности. Данные методы немного хуже по точности в сравнении с нестатистическими. Однако они отвязаны от конкретной предметной области.
4. Работы посвященные нестатистическим методам проигнорировали извлечение неявных ПСС. Тогда как машинные методы рассмотрели извлечение ПСС именно со стороны неявных и двусмысленных связей.
5. Для обучения машинных методов используются корпуса с огромным охватом: Википедия, WordNet, VerbNet, FrameNet, TREC и т.д. После такого обучения модели становятся переобученными и использовать такие модели в специализированных областях не имеет смысла. Точность в таких областях так же может упасть в связи с отсутствием достаточной информации для обучения.

6. В статьях не приводятся сравнения с существующими методами. Несмотря на то, что указаны цифры по некоторым метрикам, между собой методы невозможно сравнить так как нет какого-либо стандартизированного набора данных. Как видно из сводной табл. 3 почти все методы и модели оценивались авторами на разных наборах данных.

7. Главным недостатком всех методов является то, что они предназначены для одного языка. Преимущественно это английский язык. Для извлечения ПСС из текста на русском языке необходимо модернизировать каждый метод.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Штанчаев Х.Б.* Нестатистические методы автоматического извлечения причинно-следственных связей из текста // Известия ЮФУ Технические науки. – 2023. – № 2. – С. 273-280.
2. *Girju R.* Automatic detection of causal relations for question answering // Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering. – 2003. – Vol. 12. – P. 76-83.
3. *Girju R., Moldovan D.* Text mining for causal relations // FLAIRS Conference. – 2002. – P. 360-364.
4. *Quinlan J.R.* C4. 5: programs for machine learning. – Elsevier, 2014.
5. *Marcu D., Echihabi A.* An unsupervised approach to recognizing discourse relations // Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. – 2002. – P. 368-375.
6. *Дауни А.Б.* Д21 Байесовские модели: пер. с англ. В.А. Яроцкого. – М.: ДМК Пресс. 2018. – 182 с.
7. *Chang D.-S., Choi K.-S.* Causal relation extraction using cue phrase and lexical pair probabilities // in Natural Language Processing– IJCNLP. – 2004. – Springer, 2004. – P. 61-70.
8. *Tapanainen P., Järvinen T.* A non-projective dependency parser // Proceedings of the fifth conference on Applied natural language processing. Association for Computational Linguistics. – 1997. – P. 64-71.
9. *Blanco E., Castell N., Moldovan D.I.* Causal relation extraction. – LREC, 2008.
10. *Sil A., Huang F., Yates A.* Extracting action and event semantics from web text // AAAI Fall Symposium: Commonsense Knowledge. – 2010.
11. *Church K.W., Hanks P.* Word association norms, mutual information, and lexicography // In Proceedings of the 27th Annual Conference of the Association of Computational Linguistics. – 1989. – P. 76-83.
12. *Gordon A.S., Bejan C.A., Sagae K.* Commonsense causal reasoning using millions of personal stories. – AAAI, 2011.
13. *Bethard S., Martin J.H.* Learning semantic links from a corpus of parallel temporal and causal relations // Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers. Association for Computational Linguistics. – 2008. – P. 177-180.
14. *Rink B., Bejan C.A., Harabagiu S.M.* Learning textual graph patterns to detect causal event relations // FLAIRS Conference. – 2010.
15. *Yan X. and Han J.* Graph-based substructure pattern mining // Data Mining, Proceedings. 2002 IEEE International Conference on. – IEEE. 2002. – P. 721-724.
16. *Sorgente A., Vettigli G., Mele F.* Automatic extraction of cause effect relations in natural language text // DART@ AI\* IA. Vol. 2013. – P. 37-48.
17. *Yang X. and Mao K.* Multi level causal relation identification using extended features // Expert Systems with Applications. – 2014. – Vol. 41, No. 16. – P. 7171-7181
18. *Pakray P., Gelbukh A.* An open domain causal relation detection from paired nominal // 13th Mexican international conference on artificial intelligence (MICAI-2014). Nature-Inspired Computation and Machine Learning. – 2014. – Vol. 8857. – P 261-271.
19. *Gurulingappa H., Rajput AM., Roberts A., Fluck J., Hofmann-Apitius M., Toldo L.* Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports // J Biomed Inform. – 2012. – Vol. 45 (5). – P. 885-892.



20. Rutherford A., Xue N. Discovering implicit discourse relations through brown cluster pair representation and coreference patterns // Proceedings of the 14th conference of the European chapter of the association for computational linguistics. Association for Computational Linguistics. – 2014. – P. 645-654.
21. Xu Y., Mou L., Li G., Chen Y., Peng H., Jin Z. Classifying relations via long short-term memory networks along shortest dependency paths // Proceedings of the 2015 conference on empirical methods A survey on extraction of causal relations in natural language processing. Association for Computational Linguistics. 2015. Lisbon. Portugal. – P. 1785-1794.

#### REFERENCES

1. Shtanchaev Kh.B. Nestatisticheskie metody avtomaticheskogo izvlecheniya prichinno-sledstvennykh svyazey iz teksta [Non-statistical methods for automatically extracting cause-and-effect relationships from text], *Izvestiya YuFU Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 2, pp. 273-280.
2. Girju R. Automatic detection of causal relations for question answering, *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering*, 2003, Vol. 12, pp. 76-83.
3. Girju R., Moldovan D. Text mining for causal relations, *FLAIRS Conference*, 2002, pp. 360-364.
4. Quinlan J.R. C4. 5: programs for machine learning. Elsevier, 2014.
5. Marcu D., Echihab A. An unsupervised approach to recognizing discourse relations, *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 2002, pp. 368-375.
6. Dauni A.B. D21 Bayesovskie modeli [D21 Bayesian models]: transl. from engl. V.A. Yarockogo. Moscow: DMK Press. 2018, 182 p.
7. Chang D.-S., Choi K.-S. Causal relation extraction using cue phrase and lexical pair probabilities, in *Natural Language Processing–IJCNLP*, 2004. Springer, 2004, pp. 61-70.
8. Tapanainen P., Järvinen T. A non-projective dependency parser, *Proceedings of the fifth conference on Applied natural language processing. Association for Computational Linguistics*, 1997, pp. 64-71.
9. Blanco E., Castell N., Moldovan D.I. Causal relation extraction. LREC, 2008.
10. Sil A., Huang F., Yates A. Extracting action and event semantics from web text, *AAAI Fall Symposium: Commonsense Knowledge*, 2010.
11. Church K.W., Hanks P. Word association norms, mutual information, and lexicography, *In Proceedings of the 27th Annual Conference of the Association of Computational Linguistics*, 1989, pp. 76-83.
12. Gordon A.S., Bejan C.A., Sagae K. Commonsense causal reasoning using millions of personal stories. AAAI, 2011.
13. Bethard S., Martin J.H. Learning semantic links from a corpus of parallel temporal and causal relations, *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers. Association for Computational Linguistics*, 2008, pp. 177-180.
14. Rink B., Bejan C.A., Harabagiu S.M. Learning textual graph patterns to detect causal event relations, *FLAIRS Conference*. – 2010.
15. Yan X. and Han J. Graph-based substructure pattern mining, *Data Mining, Proceedings. 2002 IEEE International Conference on*. IEEE. 2002, pp. 721-724.
16. Sorgente A., Vettigli G., Mele F. Automatic extraction of cause effect relations in natural language text, *DART@ AI\* IA*. Vol. 2013, pp. 37-48.
17. Yang X. and Mao K. Multi level causal relation identification using extended features, *Expert Systems with Applications*, 2014, Vol. 41, No. 16, pp. 7171-7181
18. Pakray P., Gelbukh A. An open domain causal relation detection from paired nominal, *13th Mexican international conference on artificial intelligence (MICAI-2014). Nature-Inspired Computation and Machine Learning*, 2014, Vol. 8857, pp. 261-271.
19. Gurulingappa H., Rajput AM., Roberts A., Fluck J., Hofmann-Apitius M., Toldo L. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports, *J Biomed Inform*, 2012, Vol. 45 (5), pp. 885-892.

20. *Rutherford A., Xue N.* Discovering implicit discourse relations through brown cluster pair representation and coreference patterns, *Proceedings of the 14th conference of the European chapter of the association for computational linguistics. Association for Computational Linguistics*, 2014, pp. 645-654.

Статью рекомендовал к опубликованию д.т.н., профессор А.Н. Целых.

**Штанчаев Хайрутин Баширович** – Дагестанский государственный технический университет; e-mail: shtanchaev.h@gmail.com; г. Махачкала, Россия; тел.: +79883081572; кафедра ПОВТиАС; к.т.н.

**Shtanchaev Khairutin Bashirovich** – Dagestan State Technical University; e-mail: shtanchaev.h@gmail.com; Makhachkala, Russia; phone: +79883081572; the department of POVTiAS; cand. of eng. sc.