

ОГЛАВЛЕНИЕ

ГИСТОГРАММНЫЙ МЕТОД ОБРАБОТКИ ИЗОБРАЖЕНИЙ СТЕНОПИСЕЙ АРХИТЕКТУРНЫХ ПАМЯТНИКОВ

П. И. Владимиров, Е. Ю. Зыков, В. В. Кугуракова 1121–1141

УВЕЛИЧЕНИЕ РОБАСТНОСТИ НЕЙРОННЫХ СЕТЕЙ ЗА СЧЕТ ГЕНЕРАЦИИ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ, ИНВАРИАНТНЫХ К АТТРИБУТАМ

М. Р. Газизов, К. А. Григорян 1142–1154

ОПРЕДЕЛЕНИЕ ДЕФЕКТОВ НА СТАЛЬНЫХ ЛИСТАХ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Р. Д. Гаскаров, А. М. Бирюков, А. Ф. Никонов, Д. В. Агниашвили,
Д. А. Хайрисламов 1155–1171

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

А. А. Елизаров, Е. В. Разинков 1172–1191

ПЕРСИСТЕНТНЫЕ ГОМОЛОГИИ: ПРИМЕНЕНИЕ К МОНИТОРИНГУ ОБРАЗОВАНИЯ ТРЕЩИН ПРИ ГИДРОРАЗРЫВЕ ПЛАСТА

К. Ю. Ерофеев, М. Т. Зиятдинов, Е. В. Мокшин 1192–1212

ПРОСТРАНСТВЕННАЯ ОРИЕНТАЦИЯ ОБЪЕКТОВ НА ОСНОВЕ ОБРАБОТКИ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ ДЛЯ ГЕНЕРАЦИИ РАСКАДРОВОК

В. В. Кугуракова, Г. Ф. Сахибгареева, А. З. Нгуен, А. М. Астафьев 1214–1238

ПРОГРАММА «ИСТОРИЯ ГЕНИАЛЬНОГО ОТКРЫТИЯ»

Р. В. Мосолов 1239–1278

V МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ДЛЯ НАУК О ЗЕМЛЕ И ПРИЛОЖЕНИЯ ДЛЯ ГЕОЛОГИИ, ГОРНОЙ ПРОМЫШЛЕННОСТИ И ЭКОНОМИКИ. ITES&MP-2019»

В. В. Наумова 1279–1300

**ОПИСАНИЕ КОНТЕКСТНО-СВОБОДНЫХ ГРАММАТИК В ФОРМАТЕ ДАННЫХ JSON
ДЛЯ ГЕНЕРАТОРОВ СИНТАКСИЧЕСКИХ АНАЛИЗАТОРОВ**

О. К. Осипов

1301–1323

**ЦИФРОВОЙ РЕПОЗИТОРИЙ "GEOLOGYSCIENCE.RU": ОТКРЫТЫЙ ДОСТУП
К НАУЧНЫМ ПУБЛИКАЦИЯМ ПО ГЕОЛОГИИ РОССИИ**

М. И. Патук, В. В. Наумова, В. С. Ерёменко

1324–1338

УДК 72.025.3, 004.93, 535-15, 620.179.1

ГИСТОГРАММНЫЙ МЕТОД ОБРАБОТКИ ИЗОБРАЖЕНИЙ СТЕНОПИСЕЙ АРХИТЕКТУРНЫХ ПАМЯТНИКОВ

П. И. Владимиров¹, Е. Ю. Зыков², В. В. Кугуракова³

^{1,3}Казанский (Приволжский) федеральный университет, Высшая школа информационных технологий и интеллектуальных систем

²Казанский (Приволжский) федеральный университет, Институт физики

¹pavelvladimirov96@gmail.com, ²Evgeniy.Zykov@kpfu.ru,

³Vlada.Kugurakova@gmail.com

Аннотация

Представлена разработка программного обеспечения, направленного на обработку изображений стенописей архитектурных памятников с целью выявления дефектов. Для дефектоскопии использован гистограммный подход – сравнение яркостных характеристик двух изображений фресок. Данный метод позволяет отслеживать состояние архитектурных памятников с небольшими затратами и минимальным вовлечением человека. Разработанная технология применяется в рамках охранных работ объекта культурного наследия острова-града Свияжск.

Ключевые слова: дефектоскопия, обработка изображений, архитектурный памятник, объект культурного наследия, картограммирование.

ВВЕДЕНИЕ

Одной из важных задач мониторинга состояния объектов культурного наследия является наблюдение за состоянием деградирующей фресковой живописи. На данный момент времени фиксация сюжетных композиций и их сохранность состоят в графическом или фотометрическом исполнении, которое не поддается автоматизации. Данный аспект серьезно затрудняет и усложняет мониторинг состояния объектов культурного наследия со сложными архитектурными расписанными элементами интерьера. Картограммирование древних фресок – это создание полных копий всех росписей исследуемого храма не в от-

дельных фрагментах, а целиком в форме разверток стен помещения, выполненных в цвете с точным учетом масштабных соотношений всех частей расписанного интерьера. Методика фиксации сюжета и состояния фресковой живописи практически не изменялась с конца XIX в. Она состоит в выполнении схематического масштабного черно-белого изображения основных элементов живописи на полупрозрачной бумаге вручную. Однако такой метод имеет ряд недостатков. Прежде всего, чертеж выполняется в единственном экземпляре на один сюжет фрески без привязки к внутренним архитектурным элементам здания. Также на одной картограмме размещаются все дефекты живописного полотна, имеющие место на тот момент времени, в который они выполнялись, что затрудняет его расшифровку. Использование фотометрии ограничивается фотофиксацией фресок. Огромное количество художественных деталей фресок на фотографии затрудняет чтение дефектов, поэтому произвести анализ состояния живописи достаточно проблематично. Таким образом, перечисленные выше методы не позволяют провести мониторинг изменений живописного полотна с целью выявления рисков и угроз состояния и выработать мероприятия по обеспечению их сохранения. Оптимальным решением данной проблемы является периодическое картограммирование фресок с универсальным систематизированным представлением их состояния во времени.

Необходимость регулярного мониторинга обязывает к некоей систематизации, однако если использовать стандартные средства, уже давно устаревшие по сравнению с современными технологиями, это может стать довольно трудоемкой работой.

Решением данной проблемы может стать создание программного обеспечения, основанного на фотофиксации и сканировании необходимых стенописей, с их последующей обработкой и анализом [1]. Такой подход позволит автоматизировать процесс мониторинга изменений живописного полотна с целью выявления рисков и угроз состояния и выработать мероприятия по обеспечению их сохранения.

ОБЗОР ЛИТЕРАТУРЫ

В последние годы область цифровой обработки изображений быстро развивается. Защита и восстановление фресок могут быть осуществлены с использованием компьютера для виртуального восстановления фресок, в [2] анализируется технология восстановления фресок Дуньхуана, основанная на выцветшем моделировании текстур. Метод реставрационных копий может решить такие проблемы, как затухание и обесцвечивание фресок, а также воспроизвести истинный облик культурных реликвий на разных этапах и обеспечить современную техническую поддержку для защиты культурных реликвий. Защита и восстановление цифровой росписи обеспечивает достаточную научную основу и среду тестирования для физической защиты и процесса восстановления фресок и минимизирует риск работ по их защите. Предлагаемый метод имеет реконструктивный характер, но не предполагает отслеживание изменений состояния красочного слоя, включая микродефекты.

Проблема мониторинга фресок мавзолеев на наличие на них пузырьков описана в [3], где предложен метод обнаружения небольших пузырьков с улучшенными терагерцовыми изображениями. Обработка производилась с помощью системы формирования изображений спектроскопии с временным спектром ТГц. Чтобы наблюдать за маленькими пузырьками в фресках, для обработки полученных изображений принят всеобъемлющий алгоритм улучшения. Результаты показывают, что предлагаемый комплексный метод может быть использован для обнаружения крошечных дефектов фресок. Предлагаемый метод требует дорогостоящего оборудования.

Другой подход – акустическая эмиссия. Это метод неразрушающего контроля, широко используемый для мониторинга структурного состояния различных материалов, таких как бетон, кирпичная кладка и камни. С использованием данной техники было проведено исследование, целью которого стали выявление эволюции повреждений в поддержке украшенных поверхностей Ренессансного комплекса “Sacri Monti di Varallo” (Пьемонт, Италия) и использование собранных данных, поступающих из “in situ” мониторинга, для защиты работ от сейсмического риска и возможных обрушений вследствие землетрясений [4].

Данный метод предполагает мониторинг ограждающих конструкций объекта, но не применим к отслеживанию оптических характеристик.

Археологические объекты ценны и требуют методов неразрушающего контроля, в [5] освещается оптическая когерентная томография (ОКТ) на основе интерферометра Майкельсона, которая представляет собой быструю бесконтактную когерентную оптическую технологию визуализации. Данный метод, по словам авторов, подходит для тестирования любой части росписи и других культурных реликвий. Он предполагает применение дорогостоящего оборудования.

В исследовательской работе [6] приведен метод обработки изображений фресок с помощью алгоритма нахождения краев и сегментации изображений. Метод предназначен для определения границ разрушения красочного слоя и нахождения макродефектов.

Все приведенные выше работы включают в себя дефектоскопию как одну из целей. В каждой описаны свои методы различной сложности и эффективности, однако все они уступают в простоте и стоимости предлагаемому нами гистограммному экспресс-анализу.

РАЗРАБАТЫВАЕМЫЙ МЕТОД

Для автоматической обработки изображения, его дальнейшего изучения и анализа на дефекты была разработана программа, выполняющая следующие основные этапы:

1. Корректировка входного изображения для удобства его изучения оператором.
2. Проективное преобразование изображения. Построение яркостных гистограмм изображения по X- и Y-координатам пикселей.
3. Корректировка гистограмм старого и нового изображений в зависимости от корреляции между ними.
4. Сравнение старого и нового изображений фрески – дефектоскопия.

ПРОЕКТИВНЫЕ ПРЕОБРАЗОВАНИЯ ИЗОБРАЖЕНИЯ

При фотографировании фресок и стенописей далеко не всегда возможно запечатлеть изображение на стене так, чтобы оно было прямо перед фотографом. В большинстве случаев изображение получается под различными углами ввиду различных архитектурных реставраций и прочих работ над сохранением объекта культуры в целостности (рис. 1).

Для корректной обработки и дефектоскопии необходимо работать только с самим изображением фрески. Это значит, что её нужно как-то «поворачивать», избавляться от углов, обрезать. Все подобное уже придумано и рассмотрено с помощью высшей математики.



Рисунок 1. Пример фотографии фрески под углом

Конкретнее, в данном случае должны быть использованы проективные преобразования (рис. 2). Они позволяют различным образом взаимодействовать с проекциями и перспективой. Такой подход часто используется при построении карт с помощью спутниковых фотографий, также он подойдет и в оговоренном выше случае. Чтобы подобное преобразование было возможно осуществить с изображением, необходимо выбрать четыре точки преобразования P_1, \dots, P_4 (см. рис. 2).

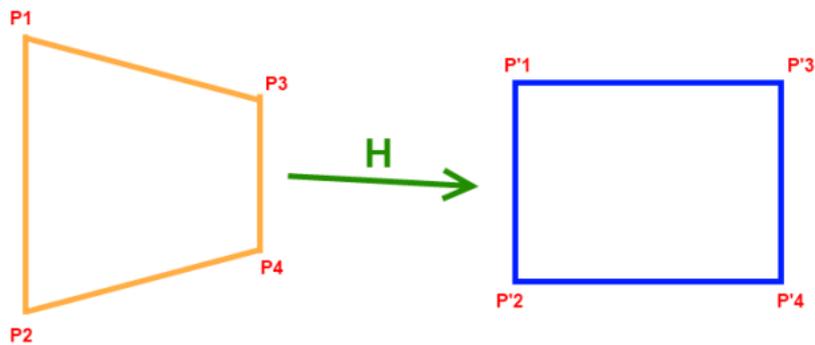


Рисунок 2. Пример проективного преобразования

Используя координаты выбранных точек и координаты четырех конечных точек (P'1, ..., P'4), с помощью специальных формул (1), (2) и матричных вычислений (3), приведенных ниже, можно получить 8 параметров, необходимых для отображения всех остальных точек (в нашем случае – пикселей изображения) из четырехугольника в прямоугольник, то есть изображение фрески можно будет «повернуть лицом» к наблюдателю.

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad (1)$$

$$y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}, \quad (2)$$

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1x'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x'_2 & -y_2x'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3x'_3 & -y_3x'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x'_4 & -y_4x'_4 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y'_2 & -y_2y'_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3y'_3 & -y_3y'_3 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y'_4 & -y_4y'_4 \end{bmatrix} \times \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \\ y'_1 \\ y'_2 \\ y'_3 \\ y'_4 \end{bmatrix}. \quad (3)$$

Для проективных преобразований был создан специальный класс матриц с необходимыми нестандартными конструкторами и методами, основным из которых является метод нахождения обратной матрицы. С её помощью находятся

8 параметров $h_{1,1} - h_{3,2}$ в (3). На сцене Unity в роли четырех точек используются обычные сферы. Зная их местоположение и используя рейкасты¹, можно найти пиксели, над которыми они находятся в данный момент, по сути, определить координаты точек преобразования. На рис. 3 можно наблюдать пример проективного преобразования над изображением фрески. Также, чтобы сферы было удобнее двигать и выбирать место на текстуре, где они будут стоять, было создано подобие увеличенного изображения текстуры. В нем показывается часть текстуры, увеличенная в несколько раз, над которой в данный момент пользователь перемещает сферу. Это создано с помощью дополнительной камеры и рендер-текстуры², помещенной в поле *Target* этой камеры (рис. 3, 4).

Однако при проективных преобразованиях изображений существуют некоторые проблемы. Одна из основных – потеря пикселей изображения после преобразования. Подобное происходит из-за того, что при отображении пикселя с целочисленными координатами далеко не факт, что новые координаты пикселя тоже будут целочисленными и будут лежать на координатной сетке нового изображения. На помощь приходит интерполяция³ пикселей. Существует несколько способов, отличающихся качеством, скоростью и сложностью реализации, в текущей версии проекта используется метод ближайшего соседа.

¹ Рейкаст (*Raycast*) – термин, который используется для обозначения луча, который может быть выпущен из какой-либо точки в пространстве и направлен в указанную сторону.

² Рендер-текстура (*RenderTexture*) – текстура, которая обновляется в реальном времени.

³ Интерполяция – способ нахождения промежуточных значений величины (в нашем случае местоположение пикселя) по имеющемуся дискретному набору известных значений.



Рисунок 3. Проективное преобразование фрески

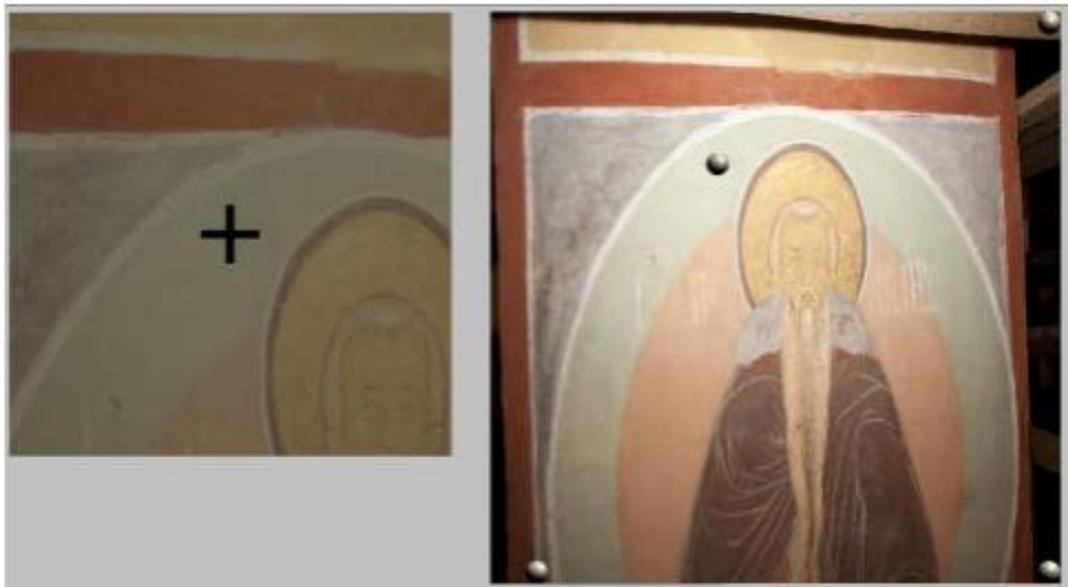


Рисунок 4. Увеличение изображения в месте точки преобразования

Подобная реализация очень простая и быстрая, есть потеря качества изображения, но на крупных изображениях, используемых в проекте, ее в принципе не заметно. Суть данного подхода заключается в следующем – после того, как пиксель был отображен в новое изображение, берется значение ближайшего пикселя, если координаты нецелочисленны.

ПОСТРОЕНИЕ ГИСТОГРАММ

Во многих случаях для обработки изображений используются гистограммы. Изучив гистограмму, можно получить общее представление о правильности экспозиции, контрасте и цветовом насыщении снимка, оценить требуемую коррекцию как при съёмке, так и при последующей обработке.

Обычно используют гистограммы для каждого цветового (красный, зелёный, синий) и яркостного каналов. Однако в нашем случае должны использоваться гистограммы не в классическом их понимании, они должны использоваться под перепады значений на определенных местах изображения, чтобы по этим перепадам было возможно найти дефектное место. Сначала строится яркостная гистограмма по всем X-координатам, по столбцам, по сути складываются значения яркостей каждого пикселя в каждом столбце, в соответствии с этим каждый столбец значений яркостей передаётся в новую текстуру. То же самое производится и по Y-координатам, строкам. На рис. 5 показаны результаты построения гистограмм данным способом.



Рисунок 5. Изображение и его гистограммы

КОРРЕЛЯЦИЯ ДЛЯ КОРРЕКТИРОВКИ И ВЫРАВНИВАНИЯ ГИСТОГРАММ

Одна часть проблемы – невозможность делать фотографии фрески всегда одинаковыми и под одним углом – решается с помощью проективных преобразований. Другая часть проблемы заключается в том, что фотографии одной и той же фрески могут быть сделаны не просто под разными углами и с разных расстояний, они могут еще и отображать разные куски настенного изображения с разных сторон. Чтобы было возможно сравнить эти изображения, необходимо на их гистограммах выделить те части, где фотографии максимально одинаковые и похожие друг на друга. Реализовать такое выделение можно с помощью подсчета линейного коэффициента корреляции⁴ между значениями в гистограммах двух разных изображений:

$$r_{XY} = \frac{\text{cov}_{XY}}{\sigma_X \sigma_Y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 (Y - \bar{Y})^2}}, \quad (4)$$

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$$

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$$

где X и Y в (4) – это значения в массивах пикселей (гистограммах) нового изображения и старого.

Общий алгоритм работы:

1. Строятся гистограммы для изображений по координатам X и Y .
2. В одной из гистограмм смещаются координаты X на некоторое число N значений влево относительно другой и вычисляется коэффициент корреляции между ними.
3. Поочерёдно изменяется смещение в диапазоне от $-N \leq \Delta x \leq N$, строится для каждого смещения коэффициент корреляции.
4. Выбирается то из смещений, коэффициент корреляции при котором

⁴ Корреляция – статистическая взаимосвязь двух или более случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми).

будет наибольшим, оно и будет искомым смещением между изображениями по координате X .

5. Изображения выравниваются по координате X , и снова строятся гистограммы, с учётом отсеечения краев изображений.

6. Пункты 2–5 повторяются для координаты Y .

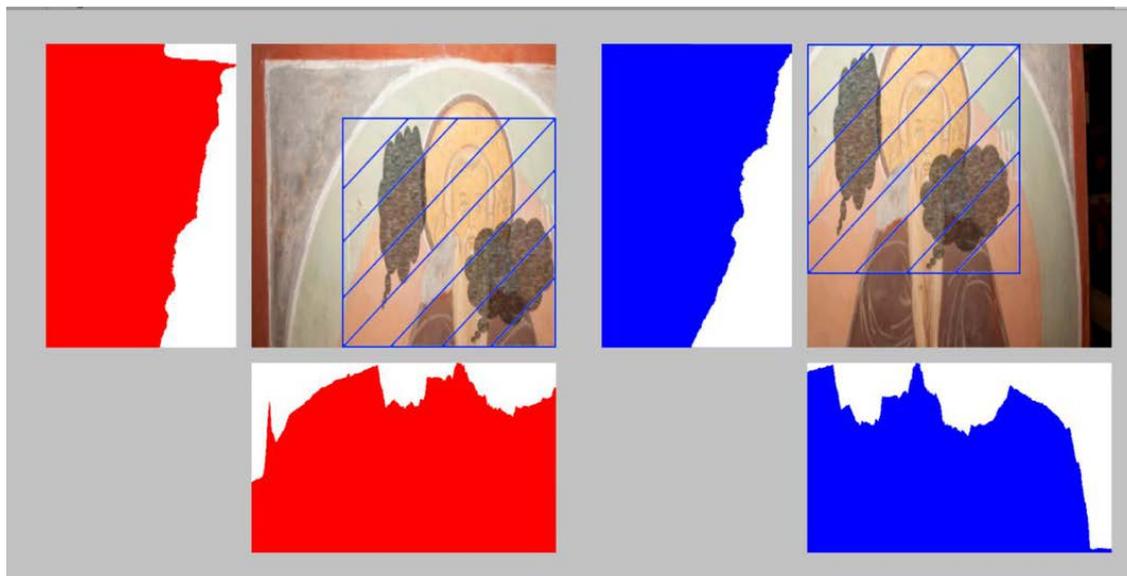


Рисунок 6. Результат работы алгоритма, синей сеткой выделены места на изображениях, которые необходимо использовать в дальнейшей работе

СРАВНЕНИЕ ИЗОБРАЖЕНИЙ И ДЕФЕКТОСКОПИЯ

После этапов загрузки изображения, его предварительной обработки и изменения наступает основной этап работы программы (рис. 7). Используя данные старого изображения фрески (т. е. сделанного некоторый промежуток времени назад) и текущего, необходимо определить, подверглось ли изображение какому-либо изменению, появились ли на нем дефекты различного характера. Для подобного анализа был использован гистограммный подход.

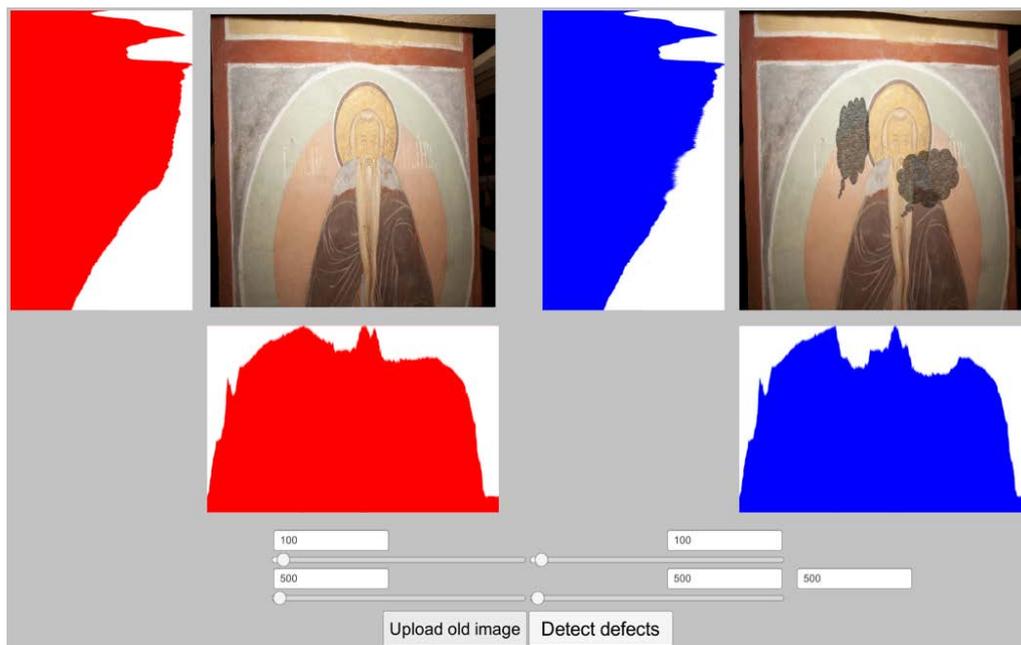


Рисунок 7. Исходное и обработанное изображения

Эти рисунки иллюстрируют, что с помощью гистограмм возможно найти на фреске дефекты – в местах, где на фреске проявились различные изменения и ухудшения, например, отколотся кусок стены либо образовалась плесень, будут затемнения, значит, яркость пикселей там будет меньше по сравнению с яркостями на прошлом изображении.

Чтобы найти дефекты, нужно пройти по гистограммам обоих изображений и выделить места, где значения различаются. Естественно, чтобы не обращать внимания на небольшие изменения и дефекты, можно использовать некоторое изменяемое пороговое значение и, если подобное отклонение присутствует, то заносить эти столбец или строку гистограммы в список дефектных. После проведения этой операции на X- и Y-гистограммах нового изображения будут известны дефектные строки и столбцы. Совместив их на фреске, получим пересечения, прямоугольники, которые и будут определять области, подвергшиеся ухудшению. Изменяя различные пороговые значения, можно изменять размеры областей, которые будут помечаться как дефектные, а также их количество (рис. 8).

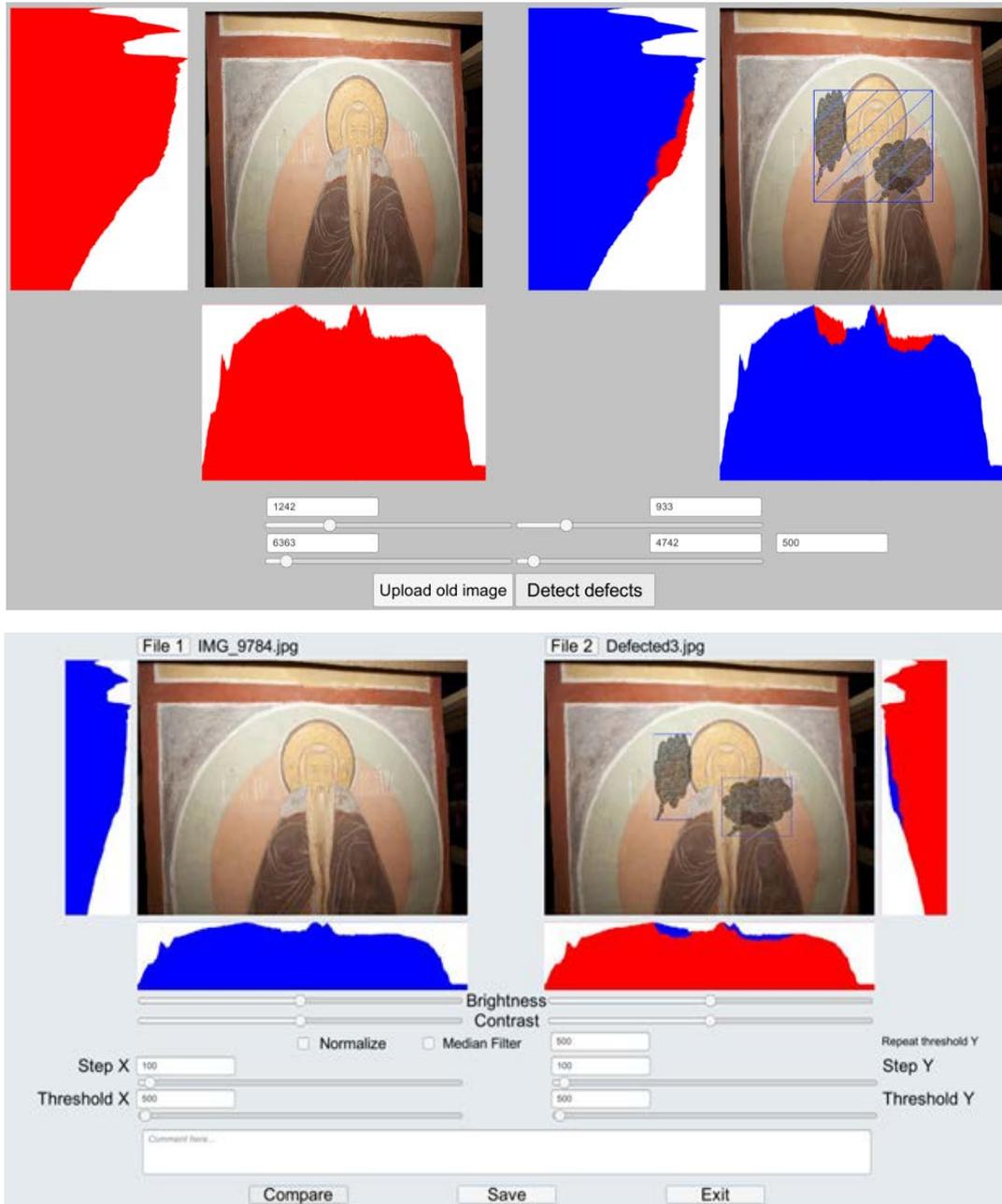


Рисунок 8. Обнаружение дефектов с различными настройками пороговых значений

УЛУЧШЕНИЕ КАЧЕСТВА ДЕФЕКТОСКОПИИ

В предыдущих примерах был продемонстрирован результат работы описанного подхода. При этом можно заметить, что «дефекты» ненастоящие и нанесены вручную в графическом редакторе для тестирования работоспособности алгоритма. Несмотря на то, что оригинальная идея в алгоритме работает

идеально на подобных ярких и больших пятнах на фотографиях, на практике почти невозможно достичь такого же результата, если не прибегнуть к дополнительным настройкам и изменениям.

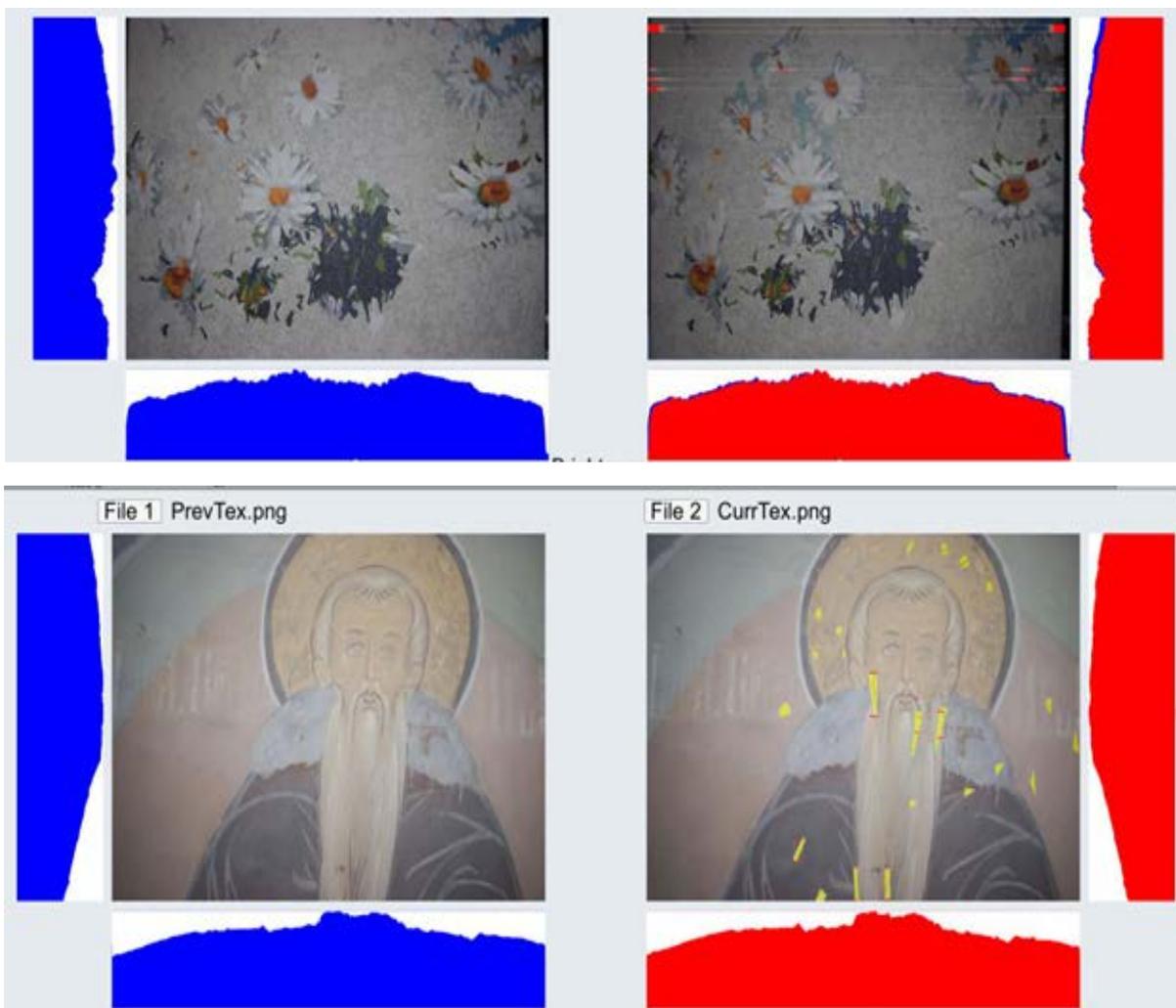


Рисунок 9. Некорректная обработка реалистичных дефектов

В реальной жизни дефекты, возникающие с течением времени на стенописях, гораздо более мелкие и тусклые по сравнению с нарисованными черными кляксами. В такой ситуации обработка изображения напрямую и полностью за один заход в большинстве случаев не принесет результата, так как пикселей на изображении слишком много, а значения их яркостей не различаются настолько сильно, чтобы можно было их корректно обработать и показать пользователю.

Рисунки 9 демонстрируют работу алгоритма с более реалистичными дефектами, точнее, ее отсутствие. На рис. 10 отсутствуют видимые различия между гистограммами.



Рисунок 10. Дефектоскопия с декомпозицией

Чтобы решить данную проблему, можно прибегнуть к способу декомпозиции, конкретно – подразбиению изображения на части. Так как значения яркостей на гистограммах несильно отличаются друг от друга в глобальном плане, то можно попытаться разделить исходное изображение на 3, 4, 6, 8 или 9 частей, увеличить их (опционально) и после этого запустить алгоритм на каждой отдельной части изображения. Таким образом найти дефекты на отдельных и увеличенных изображениях будет гораздо проще, после чего их координаты можно будет сохранить и позже отобразить на исходном изображении (см. рис. 10).

С помощью данного подхода дефектоскопия изображений производится на порядок лучше, появляется возможность найти больше мелких дефектов, но стоит заметить, что не все из них будут нам важны и необходимы, а некоторые дефекты остаются все еще сложно «найти». Исходя из этого, алгоритм нуждается в дальнейшей доработке и применении, так как имеет наивысший потенциал среди перечисленных.

В результате было выполнено большинство основных задач в решении проблемы автоматизации обработки стенописей. Другими шагами являются высокочастотная фильтрация, масштабирование изображения, сравнение гистограмм изображений разного размера, полная автоматизация дефектоскопии.

ЗАКЛЮЧЕНИЕ

Высококачественные пространственные данные высокой плотности, которые способна захватить пара инфракрасных различных новейших камер (использующих, например, технологию Google Project Tango), можно использовать для трёхмерного сканирования помещений, построения карт глубины и реконструкции захваченных сцен. Широкоугольная камера, гиростабилизатор и выверенные программные механизмы синхронизации сенсоров через фиксирование отметок времени позволяют таким устройствам использовать эти возможности, в частности, для быстрой двухдиапазонной 3D-дефектоскопии стенописей архитектурных памятников, автоматизируя процесс мониторинга сложных текстурированных объектов с существенной точностью фиксирования плоских изображений на внутренних поверхностях объемных архитектурных элементов.

Предложенная технология экспресс-оценки состояния стенописей разрабатывается как модуль системы мониторинга состояния архитектурных памятников [10; 11], выполняется в рамках работ по охранному мероприятию объекта культурного наследия федерального уровня острова-града Свяжск.

По материалам проведенной работы зарегистрировано свидетельство на программу для ЭВМ [12].

СПИСОК ЛИТЕРАТУРЫ

1. *Sitdikov A.G., Shaykhutdinova E.F., Kugurakova V., Zykov E.Yu., Kasimov A.V., Skobelev A.A.* Development of Modern Methods for the Diagnostics of Murals in Architectural Monuments // *European Research Studies Journal*. 2017. Vol. XX, Special Issue. P. 496–506.
2. *Liang Y.* Research on restoration technology of dunhuang frescoes based on faded texture simulation// *Boletin Tecnico/Technical Bulletin*. 2017. Vol. 55, No. 13. P. 737–743.
3. *Yuan M., Sun W., Wang X., Ye J., Wang S., Zhang Q., Zhang Y.* Mural defects detection with enhanced terahertz images// *Journal of Electronic Science and Technology*. 2016. Vol. 14, No. 4. P. 337–342.
4. *Accornero F., Invernizzi S., Lacidogna G., Carpinteri A.* The sacred mountain of varallo renaissance complex in Italy: Damage analysis of decorated surfaces and structural supports// *Acoustic, Electromagnetic, Neutron Emissions from Fracture and Earthquakes*. 2016. P. 249–264.
5. *Yang S., Zhu R., Mi L., Cao Y., Li Q.* Application of optical coherence tomography in the detection of the mural// *Guangxue Xuebao/Acta Optica Sinica*. 2015. Vol. 35, No. 5.
6. *Ma Y.* Application of edge detection and image segmentation algorithm of image processing in murals copy// *Revista de la Facultad de Ingenieria*. 2017. Vol. 32, No. 5. P. 809–818.
7. *Кугуракова В.В., Зыков Е.Ю., Касимов А.В., Ситдииков А.Г., Скобелев А.А., Шайхутдинова Е.Ф.* In situ двухдиапазонная 3D-дефектоскопия стенописей архитектурных памятников // *Электронные библиотеки*. 2016. Т. 19, №6. С. 538–558.
8. *Зыков Е.Ю., Шерстюков О.Н., Акчурин А.Д.* Исследовательский ионизонд «Циклон» Казанского университета и программное обеспечение автоматической обработки ионограмм // *Гелиогеофизические исследования*. 2013. Вып. 4. С. 39–46.
9. *Зыков Е.Ю., Акчурин А.Д., Сапаев А.Л., Шерстюков О.Н.* Автоматическая интерпретация ионограмм вертикального зондирования // *Ученые записки Казанского университета*. 2008. Т. 150, Кн. 3. С. 36–45.

10. *Sitdikov A.G., Shaykhutdinova E.F., Kasimov A.V., Azizov T.R., Kugurakova V.V.* A practical approach to an automated accounting system development of the study results of historical and architectural sites of the island of Sviyazhsk // *Helix*. 2018. Vol. 8(1). P. 2301–2306.

11. *Azizov T.R., Shaykhutdinova E.F., Sitdikov A.G.* Practical Approach to Development Automated System of Record Research Results Architectural Cultural Heritage Objects Island-Town Sviyazhsk // *World Academy of Science, Engineering and Technology, International Science Index, Computer and Information Engineering*. 2015. V. 2(12). P. 1715.

12. Свидетельство о государственной регистрации программы для ЭВМ № 201861634 «Гистограммный анализатор изображений для дефектоскопии фресковой живописи» / Кугуракова В.В., Владимиров П.И., Зыков Е.Ю., Ситди-ков А.Г., Шайхутдинова Е.Ф. № 201861634; заявл. 19.04.2018; в реестре программ для ЭВМ с 01.06.2018.

THE HISTOGRAM APPROACH FOR COMPARING CARTOGRAMS OF MURALS

P. I. Vladimirov, E. Yu. Zykov, V. V. Kugurakova

^{1–3} *Kazan Federal University*

¹pavelvladimirov96@gmail.com, ²Evgeniy.Zykov@kpfu.ru,

³Vlada.Kugurakova@gmail.com

Abstract

The article describes the development of software aimed at processing images of murals of architectural monuments in order to identify defects. For flaw detection, a histogram approach was used – a comparison of the brightness characteristics of two images of frescoes. This method allows you to track the status of architectural monuments with low costs and minimal human involvement. The developed technol-

ogy is used as part of the protection of the cultural heritage of the island-town of Sviyazhsk.

Keywords: *defects detection, image processing, architectural monument, cultural heritage, mapping.*

REFERENCES

1. *Sitdikov A.G., Shaykhutdinova E.F., Kugurakova V., Zykov E.Yu., Kasimov A.V., Skobelev A.A.* Development of Modern Methods for the Diagnostics of Murals in Architectural Monuments // *European Research Studies Journal*. 2017. Vo. XX, Special Issue. P. 496–506.
2. *Liang Y.* Research on restoration technology of dunhuang frescoes based on faded texture simulation// *Boletin Tecnico/Technical Bulletin*. 2017. Vol. 55, No. 13. P. 737–743.
3. *Yuan M., Sun W., Wang X., Ye J., Wang S., Zhang Q., Zhang Y.* Mural defects detection with enhanced terahertz images// *Journal of Electronic Science and Technology*. 2016. Vol. 14, No. 4. P. 337–342.
4. *Accornero F., Invernizzi S., Lacidogna G., Carpinteri A.* The sacred mountain of varallo renaissance complex in Italy: Damage analysis of decorated surfaces and structural supports// *Acoustic, Electromagnetic, Neutron Emissions from Fracture and Earthquakes*. 2016. P. 249–264.
5. *Yang S., Zhu R., Mi L., Cao Y., Li Q.* Application of optical coherence tomography in the detection of the mural// *Guangxue Xuebao/Acta Optica Sinica*. 2015. Vol. 35, No. 5.
6. *Ma Y.* Application of edge detection and image segmentation algorithm of image processing in murals copy// *Revista de la Facultad de Ingenieria*. 2017. Vol. 32, No. 5. P. 809–818.
7. *Kugurakova V.V., Zykov E. Ju., Kasimov A.V., Sitdikov A.G., Skobelev A.A., Shajhutdinova E.F.* In situ dual band 3D flaw detection of murals or architectural monuments // *Jelektronnye biblioteki*. 2016. T. 19. №6. S. 538–558 (In Russian).
8. *Zykov E.Ju., Sherstjukov O.N., Akchurin A.D.* Kazan University Research “Ciklon” Iono-sonde and Automatic Ionogram Processing Software // *Geliogeofizicheskie issledovanija*. 2013. T. 4. S. 39–46 (In Russian).

9. Zykov E.Ju., Akchurin A.D., Sapaev A.L., Sherstjukov O.N. Automatic interpretation of vertical sounding ionograms // *Uchenye zapiski KGU*. 2008. T. 150 (3). S. 36–45 (In Russian).

10. Sitdikov A.G., Shaykhutdinova E.F., Kasimov A.V., Azizov T.R., Kugurakova V.V. A practical approach to an automated accounting system development of the study results of historical and architectural sites of the island of Sviyazhsk // *Helix*. 2018. Vol. 8(1). P. 2301–2306.

11. Azizov T.R., Shaykhutdinova E.F., Sitdikov A.G. Practical Approach to Development Automated System of Record Research Results Architectural Cultural Heritage Objects Island-Town Sviyazhsk // *World Academy of Science, Engineering and Technology, International Science Index, Computer and Information Engineering*. 2015. V. 2(12). P. 1715.

12. Certificate of state registration of a computer program № 201861634 «Histogram analyzer of images for defects detection of mural paintings» / Kugurakova V.V., Vladimirov P.I., Zykov E.Ju., Sitdikov A.G., Shaykhutdinova E.F. № 201861634; stat. 19.04.2018; in the registry of computer programs from the 01.06.2018.

СВЕДЕНИЯ ОБ АВТОРАХ



ВЛАДИМИРОВ Павел Игоревич – магистр ИТИС КФУ.

Сфера научных интересов: программирование на Unity.

Pavel Igorevich VLADIMIROV – master of the Higher Institute of Information Technology and Intelligent Systems.

Area of interests: Unity development.

email: pavelvladimirov96@gmail.com



ЗЫКОВ Евгений Юрьевич – к. ф.-м. н., доцент Института физики КФУ. Квалификация: радиофизик, научные интересы: распространение радиоволн, ионосфера, а также алгоритмы искусственного интеллекта и распознавания образов.

Evgeniy Yurievich ZYKOV is currently Associate Professor in the Institute of Physics of the KFU. He received Ph.D. degree in Physics and Mathematics. His qualification radio physicist, research interests include the radiowave propagation, ionosphere, and also artificial intelligence and pattern-recognition algorithms.

email: evgeniy.zykov@kpfu.ru



КУГУРАКОВА Влада Владимировна – к. т. н., доцент кафедры программной инженерии ИТИС КФУ, руководитель НИЛ разработки AR/VR приложений и компьютерных игр. Сфера научных интересов – генерация реалистичной визуализации трехмерных сцен и различные аспекты проектирования игр.

Vlada Vladimirovna KUGURAKOVA, PhD., Docent of ITIS KFU, Head of Laboratory «AR/VR applications and Gamedev». Research interests include different approaches in game design.

email: vlada.kugurakova@gmail.com

Материал поступил в редакцию 30 октября 2020 года

УДК 004

УВЕЛИЧЕНИЕ РОБАСТНОСТИ НЕЙРОННЫХ СЕТЕЙ ЗА СЧЕТ ГЕНЕРАЦИИ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ, ИНВАРИАНТНЫХ К АТТРИБУТАМ

М. Р. Газизов¹, К. А. Григорян²

Казанский (Приволжский) федеральный университет, Казань

¹gazizovmarat@gmail.com, ²karigri@yandex.ru

Аннотация

Робастность модели к незначительным отклонениям в распределении исходных данных является важным критерием во многих задачах. Нейронные сети могут показывать высокую точность (accuracy) на обучающей выборке, но при этом качество на тестовой выборке может сильно падать из-за разного распределения данных, причем ситуация только усугубляется на уровне подгрупп внутри каждой категории.

В данной статье мы показываем, как робастность модели на уровне подгрупп может быть значительно улучшена с помощью подхода, основанного на доменной адаптации векторных представлений. Мы обнаружили, что применение состязательного подхода к ограничению векторных представлений дает существенный прирост метрики точности (accuracy) в сложной подгруппе по сравнению с предыдущими моделями. Метод протестирован на двух независимых наборах данных, точность в сложной подгруппе на наборе данных Waterbirds составляет 90.3 {y : waterbirds; a : landbackground}, а на наборе данных CelebA – 92.22 {y : blondhair; a : male}.

Ключевые слова: робастная классификация, классификация изображений, генеративно-состязатель сети, доменная адаптация.

ВВЕДЕНИЕ

В процессе обучения нейронных сетей в общем случае пытаются минимизировать среднее значение целевой функции ошибки на обучающей выборке или, другими словами, минимизируют эмпирический риск [1]. При этом мини-

мизация средней оценки целевой функции справедлива в случаях, когда тестовая выборка получена из распределения, независимого или идентичного обучающей выборке. В подобных задачах современные модели показывают высокое качество [2], но при этом проваливаются на примерах, редких и нетипичных для обучающей выборки [3–5]. Часто это делает невозможным применение моделей в реальной жизни, если эти модели используются для принятия решений [6] или оценки рисков [7].

К примеру, модель может полагаться на ложные корреляции атрибутов изображений с целевой категорией из-за неравномерного распределения значений атрибутов внутри категории, как в случае с метками на коже (атрибут) и наличием меланомы (целевая категория) у человек. На рисунке 1 показано неравномерное распределение меток между категориями злокачественных и доброкачественных образований, т. е. метки встречаются от 4.5 до 6 раз реже в категории со злокачественным образованием, чем в других категориях. Это приводит к увеличению доли ложно-отрицательных прогнозов в группе людей со злокачественными образованиями и метками на коже [8]. В этом и подобных случаях средняя оценка целевой функции не будет достоверно отображать устойчивость модели к наличию данных меток, а качество в сложной подгруппе может значительно отличаться в меньшую сторону от среднего по всей категории.

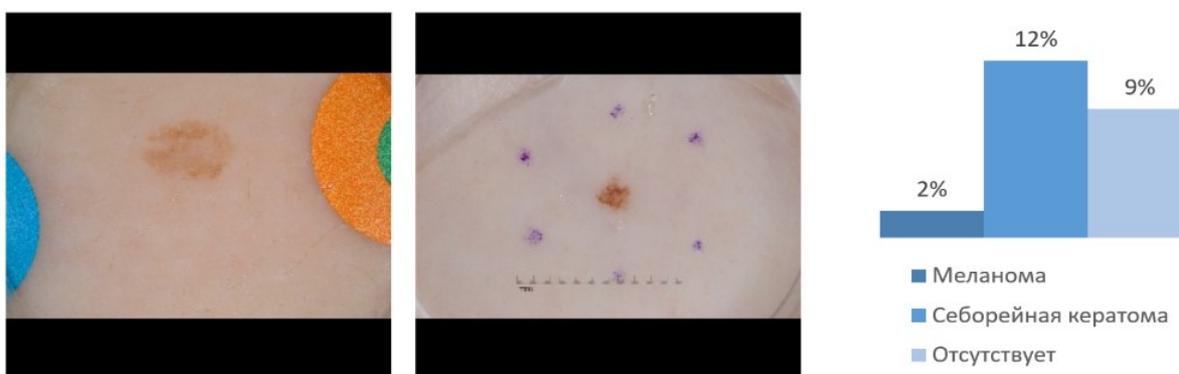


Рисунок. 1. Примеры изображений из набора данных ISIC [9–11] с метками и распределение меток среди категорий. Меланома – злокачественное образование. Себорейная кератома – доброкачественное образование. Последний столбец указывает на отсутствие образований.

Таким образом, требуется следить не только за средней оценкой показателя качества на уровне категории, но также замерять качество на уровне каждой из подгрупп внутри категории.

В данной статье мы изучаем подход, основанный на генерации инвариантных к атрибутам векторных представлений с помощью состязательных сетей на примере задачи классификации изображений. Мы предлагаем использовать доменную адаптацию для сближения векторных представлений примеров из разных подгрупп так, чтобы они становились неотличимы друг от друга на уровне категории, но все еще отличались на уровне целевого класса. Мы показываем, как адаптация векторных представлений может быть применима для минимизации зазора показателей качества между подгруппами, что значительно увеличивает устойчивость модели во время применения на тестовых наборах данных.

Предложенное нами решение показывает большую точность (ассурасу) в сложных подгруппах на двух модельных наборах данных по сравнению с другими методами. Достигнутая нами точность в сложной подгруппе для набора данных Waterbirds составляет 90.3 { $y : waterbirds; a : landbackground$ }, а для набора данных CelebA – 92.22 { $y : blondhair; a : male$ }.

ОБЗОР ЛИТЕРАТУРЫ

Group distributionally robust optimization – GDRO [12]

Задача увеличения робастности модели может быть рассмотрена на уровне процесса оптимизации. В стандартном подходе весовые коэффициенты для разных подгрупп являются одинаковыми. Мы можем влиять на них явно через задание разных весов для разных групп или неявно через семплирование данных. Авторы этой статьи предлагают подход, основанный на динамической оптимизации весовых коэффициентов для каждой подгруппы во время процесса оптимизации модели. Также ими обнаружено, что увеличение регуляризации крайне важно для увеличения точности в сложной подгруппе. Они показывают теоретическую сходимость задачи оптимизации на выпуклых задачах. Помимо этого, ими предложен новый набор данных Waterbirds для сравнения качества работы разных моделей.

Class-conditional Learned Augmentations for Model Patching – CLAMP [13]

Задача доменной адаптации может решаться как на уровне векторных представлений, так и на уровне входных данных. Авторы данного подхода предлагают использовать генеративно-состязательные сети для трансформации исходных данных между подгруппами с сохранением целевого класса. Они исследуют подход, основанный на архитектуре нейронной сети CycleGAN [14]. Такая сеть позволяет решать задачу image-to-image translation без наличия спаренных изображений, в отличие от большинства методов, которые требуют наличия пар изображений из разных доменов [15, 16].

Для увеличения консистентности генерируемых изображений авторы предлагают использовать дополнительный модуль, задача которого – приближение синтетических данных к реальным данным.

Из положительных сторон генеративного подхода можно выделить его обобщающую способность, т. к. при наличии большого объема данных и достаточных по емкости моделей можно создавать данные, максимально близкие к данным из реального распределения [17, 18], тем самым устраняя дисбаланс между подгруппами.

Из минусов данного подхода можно выделить сложность обучения генеративных моделей [19] и высокие требования к вычислительным мощностям [18].

МЕТОДОЛОГИЯ

В данном разделе мы описываем предлагаемый состязательный подход к увеличению робастности модели (adversarial approach to increase model robustness – AAIMR).

Пусть дана выборка $x \in X$, целевая переменная $y \in Y$ и атрибут $a \in A$. Наша цель – обучить классификатор $f: C(F(x)) \rightarrow y$, который будет устойчив к изменению значения атрибута $a \in A$. При этом значение атрибута a известно только при обучении модели, но неизвестно во время ее применения. Таким образом, алгоритм не должен напрямую зависеть от a .

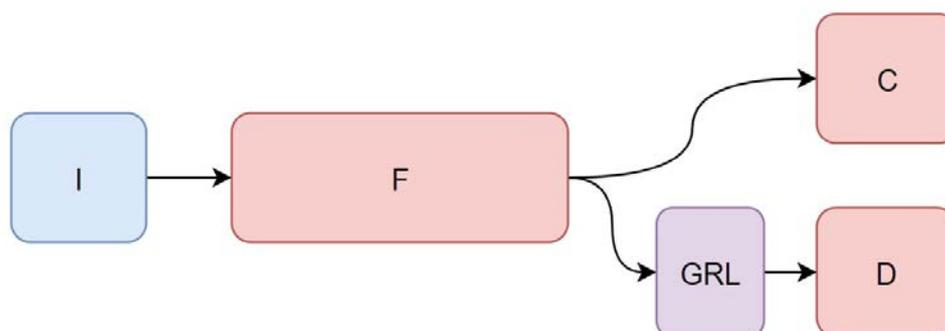


Рисунок 2. F – модель, преобразующая изображение в некоторое векторное представление, C – голова для предсказания значений целевого класса, D – дискриминатор для предсказания значений атрибута, GRL – слой обратных градиентов.

Очевидно, что переобучение модели C на значениях атрибутов a возможно только в том случае, если вектор $e = F(x)$ содержит информацию, позволяющую обучить такой классификатор $f: D(e) \rightarrow a$. Следовательно, если информация об атрибуте a не может быть получена из вектора e , то модель C будет инвариантна к значению атрибута a .

Таким образом, мы хотим поставить состязательную минимакс-задачу следующим образом: D обучается распознавать значение атрибута a , C – значение атрибута y . Задача F – научиться генерировать такие векторные представления e , чтобы дискриминатор D не мог правильно восстановить значение атрибута исходного изображения, при этом C мог правильно восстановить значение целевого класса. Для этого требуется оптимизировать следующий функционал:

$$E(\theta_f, \theta_c, \theta_d) = L(C(F(x|\theta_f)|\theta_c), y) - \lambda L(D(F(x|\theta_f)|\theta_d), a), \quad (1)$$

$$(\hat{\theta}_f, \hat{\theta}_c) = E(\theta_f, \theta_c, \hat{\theta}_d) \xrightarrow{\theta_f, \theta_c} \min, \quad \hat{\theta}_d = E(\hat{\theta}_f, \hat{\theta}_c, \theta_d) \xrightarrow{\theta_d} \max,$$

где L – целевая функция (в нашем случае – перекрестная энтропия), $\theta_f, \theta_c, \theta_d$ – оптимизируемые параметры моделей, λ – весовой коэффициент.

Чтобы функционал (1) можно было оптимизировать градиентным спуском с помощью метода обратного распространения ошибки, мы добавляем в архитектуру начальной сети слой обратных градиентов [20], как показано на рисунке 2.

В случае сходимости модель F будет продуцировать векторные представления, инвариантные к значению атрибута a , и, следовательно, модель C также будет инвариантна к изменению атрибута a .

Для обучения была взята модель ResNet-50 [21] с предобученными весами на наборе данных ImageNet [22]. В качестве оптимизатора использовался Adam [23]. Политика изменения скорости обучения не использовалась, скорость обучения была константной на всем протяжении процесса оптимизации.

Модель и другие параметры были выбраны в соответствие с [12] для возможности прямого сравнения результатов.

Для задачи Waterbirds [12] параметры оптимизации были зафиксированы $\text{weights decay} = 1.0$ и $\text{learning rate} = 0.00001$, для CelebA [20] $\text{weights decay} = 0.1$ и $\text{learning rate} = 0.00001$. Размер батча был зафиксирован на значении 128 для обеих моделей. Примеры были равномерно взвешены на уровне подгрупп. Весовой коэффициент слоя обратных градиентов $\lambda=0.5$.

Мы также дополнительно исследовали верхнюю оценку точности на задаче Waterbirds [12], для этого использовали маски, размеченные вручную. По маске вырезали объекты и зануляли фон, дальше обучали модель ResNet-50 [21] с равномерным взвешиванием примеров на уровне категорий, остальные параметры остались без изменений.

РЕЗУЛЬТАТЫ

Результаты, представленные ниже, получены на тестовых наборах данных для лучшей модели. В качестве таковой бралась та, которая имеет точность на всем валидационном наборе данных выше 0.9 и максимальную точность в сложной подгруппе.

Результаты, представленные в Таблице 1 для двух модельных задач, показывают, что предложенное решение дает значимое увеличение метрики качества в сложной группе при сохранении общей точности на том же уровне по сравнению с предыдущими подходами. Помимо этого, зазор между робастной точностью и общей точностью минимален.

Таблица 1. Результаты сравнения моделей. Общая точность считается на всей тестовой выборке. Робастная точность – это точность (accuracy) на сложной подгруппе

	CelebA		Waterbirds	
	Общая точность	Робастная точность	Общая точность	Робастная точность
ERM	94.8	41.1	97.3	60
GDRO [12]	91.8	88.3	93.2	86
CLAMP [13]	92.9	83.9	90.89	89.12
AAIMR	91.74	92.22	90.51	90.03

Таблица 2: Верхняя робастная оценка для набора данных Waterbirds. Общая точность считается на всей тестовой выборке. Робастная точность – это точность (accuracy) на сложной подгруппе

	Waterbirds	
	Средняя оценка	Робастная оценка
ERM	90.29	89.25

Результаты верхней оценки модели классификации на наборе данных Waterbirds показывают, что в данной задаче с большой вероятностью прирост точности за счет устранения корреляции между атрибутами и целевой переменной исчерпывается на уровне, указанном в Таблице 2, и дальнейшее улучшение связано с другими факторами.

ОБСУЖДЕНИЕ

Мы проанализировали подход доменной адаптации применительно к адаптации атрибутов изображений. Инвариантные к атрибутам признаки дают существенный прирост точности модели в сложной подгруппе и, как следствие, увеличивают робастность модели.

Наши результаты показывают, что доменная адаптация – перспективное направление для улучшения робастности модели.

В дальнейшем более глубокий анализ методов доменной адаптации, трансформации изображений между доменами и разделения векторных пространств может открыть немало интересных подходов к повышению робастности моделей.

СПИСОК ЛИТЕРАТУРЫ

1. *Vladimir Vapnik*. Principles of risk minimization for learning theory // Advances in Neural Information Processing Systems. 1992. P. 831–838.
2. *Christian Szegedy et al.* Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-first AAAI Conference on Artificial Intelligence, 2017.
3. *Dirk Hovy, Anders Søgaard*. Tagging performance correlates with author age // Proceedings of the 53rd annual meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (volume 2: Short papers). 2015, P. 483–488.
4. *Nicole Shadown*. Ethics and bias in machine learning: A technical study of what makes us “good”. The Transhumanism Handbook. Springer, 2019. P. 247–261.
5. *Osonde A Osoba, William Welser IV*. An intelligence in our image: The risks of bias and errors in artificial intelligence. Rand Corporation, 2017.
6. *Shai Danziger, Jonathan Levav, Liora Avnaim-Pesso*. Extraneous factors in judicial decisions // Proceedings of the National Academy of Sciences 108.17 (2011). P. 6889–6892.
7. *Amitabha Mukerjee et al.* Multi-objective evolutionary algorithms for the risk-return trade-off in bank loan management // International Transactions in Operational Research 9.5. 2002. P. 583–597.
8. *Julia K. Winkler et al.* Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition // JAMA Dermatology 155.10. 2019. P. 1135–1141.
9. *Philipp Tschandl, Cliff Rosendahl, Harald Kittler*. The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions // Scientific Data 5. 2018. P. 180161.

10. *Noel CF Codella at all.* Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC)// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE. 2018. P. 168–172.
11. *Marc Combalia at all.* Bcn20000: Dermoscopic lesions in the wild // arXiv preprint arXiv:1908.02288, 2019.
12. *Shiori Sagawa at all.* Distributionally Robust Neural Networks for Group Shifts: On the Importance of Regularization for Worst-Case Generalization // arXiv preprint arXiv:1911.08731, 2019.
13. *Sharon Li Karan Goel Albert Gu, Chris R'e.* Automating the Art of Data Augmentation CLAMP: An Instantiation of Model Patching, 2020.
URL: <http://hazyresearch.stanford.edu/data-aug-part-4>.
14. *Jun-Yan Zhu at all.* Unpaired image-to-image translation using cycle-consistent adversarial networks // Proceedings of the IEEE international Conference on Computer Vision, 2017. P. 2223–2232.
15. *Phillip Isola at all.* Image-to-image translation with conditional adversarial networks // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. P. 1125–1134.
16. *Soumya Tripathy, Juho Kannala, Esa Rahtu.* Learning image-to-image translation using paired and unpaired training samples // Asian Conference on Computer Vision. Springer, 2018. P. 51–66.
17. *Ivan Anokhin at all.* High-Resolution Daytime Translation Without Domain Labels // arXiv preprint arXiv:2003.08791, 2020.
18. *Tero Karras at all.* Analyzing and improving the image quality of stylegan // arXiv preprint arXiv:1912.04958, 2019.
19. *Sangwoo Mo, Minsu Cho, Jinwoo Shin.* Instagan: Instance-aware imageto-image translation // arXiv preprint arXiv:1812.10889, 2018.
20. *Yaroslav Ganin, Victor Lempitsky.* Unsupervised domain adaptation by backpropagation // arXiv preprint arXiv:1409.7495, 2014.
21. *Ying Tai, Jian Yang, Xiaoming Liu.* Image super-resolution via deep recursive residual network // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. P. 3147–3155.

22. Jia Deng et al. Imagenet: A large-scale hierarchical image database// 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE. 2009. P. 248–255.

23. Diederik P Kingma, Jimmy Ba. Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980, 2014.

OF NEURAL NETWORK MODEL ROBUSTNESS THROUGH GENERATING INVARIANT TO ATTRIBUTES EMBEDDINGS

Marat Gazizov¹, Karen Grigoryan²

Kazan (Volga region) Federal University, Kazan

¹gazizovmarat@gmail.com, ²karigri@yandex.ru

Abstract

Model robustness to minor deviations in the distribution of input data is an important criterion in many tasks. Neural networks show high accuracy on training samples, but the quality on test samples can be dropped dramatically due to different data distributions, a situation that is exacerbated at the subgroup level within each category. In this article we show how the robustness of the model at the subgroup level can be significantly improved with the help of the domain adaptation approach to image embeddings. We have found that application of a competitive approach to embeddings limitation gives a significant increase of accuracy metrics in a complex subgroup in comparison with the previous models. The method was tested on two independent datasets, the accuracy in a complex subgroup on the Waterbirds dataset is 90.3 {y : waterbirds;a : landbackground}, on the CelebA dataset is 92.22 {y : blondhair;a : male}.

Keywords: robust classification, image classification, generative adversarial networks, domain adaptation

REFERENCES

1. *Vladimir Vapnik*. Principles of risk minimization for learning theory // Advances in Neural Information Processing Systems. 1992. P. 831–838.
2. *Christian Szegedy et al.* Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-first AAAI Conference on Artificial Intelligence, 2017.
3. *Dirk Hovy, Anders Søgaard*. Tagging performance correlates with author age // Proceedings of the 53rd annual meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (volume 2: Short papers). 2015, P. 483–488.
4. *Nicole Shadownen*. Ethics and bias in machine learning: A technical study of what makes us “good”. The Transhumanism Handbook. Springer, 2019. P. 247–261.
5. *Osonde A Osoba, William Welser IV*. An intelligence in our image: The risks of bias and errors in artificial intelligence. Rand Corporation, 2017.
6. *Shai Danziger, Jonathan Levav, Liora Avnaim-Pesso*. Extraneous factors in judicial decisions // Proceedings of the National Academy of Sciences 108.17 (2011). P. 6889–6892.
7. *Amitabha Mukerjee et al.* Multi-objective evolutionary algorithms for the risk-return trade-off in bank loan management // International Transactions in Operational Research 9.5. 2002. P. 583–597.
8. *Julia K. Winkler et al.* Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition // JAMA Dermatology 155.10. 2019. P. 1135–1141.
9. *Philipp Tschandl, Cliff Rosendahl, Harald Kittler*. The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions // Scientific Data 5. 2018. P. 180161.
10. *Noel CF Codella et al.* Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC)// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE. 2018. P. 168–172.

11. *Marc Combalia at all.* Bcn20000: Dermoscopic lesions in the wild // arXiv preprint arXiv:1908.02288, 2019.
12. *Shiori Sagawa at all.* Distributionally Robust Neural Networks for Group Shifts: On the Importance of Regularization for Worst-Case Generalization // arXiv preprint arXiv:1911.08731, 2019.
13. *Sharon Li Karan Goel Albert Gu, Chris R'e.* Automating the Art of Data Augmentation CLAMP: An Instantiation of Model Patching, 2020. URL: <http://hazyresearch.stanford.edu/data-aug-part-4>.
14. *Jun-Yan Zhu at all.* Unpaired image-to-image translation using cycle-consistent adversarial networks // Proceedings of the IEEE international Conference on Computer Vision, 2017. P. 2223–2232.
15. *Phillip Isola at all.* Image-to-image translation with conditional adversarial networks // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. P. 1125–1134.
16. *Soumya Tripathy, Juho Kannala, Esa Rahtu.* Learning image-to-image translation using paired and unpaired training samples // Asian Conference on Computer Vision. Springer, 2018. P. 51–66.
17. *Ivan Anokhin at all.* High-Resolution Daytime Translation Without Domain Labels // arXiv preprint arXiv:2003.08791, 2020.
18. *Tero Karras at all.* Analyzing and improving the image quality of stylegan // arXiv preprint arXiv:1912.04958, 2019.
19. *Sangwoo Mo, Minsu Cho, Jinwoo Shin.* Instagan: Instance-aware imageto-image translation // arXiv preprint arXiv:1812.10889, 2018.
20. *Yaroslav Ganin, Victor Lempitsky.* Unsupervised domain adaptation by backpropagation // arXiv preprint arXiv:1409.7495, 2014.
21. *Ying Tai, Jian Yang, Xiaoming Liu.* Image super-resolution via deep recursive residual network // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. P. 3147–3155.
22. *Jia Deng at all.* Imagenet: A large-scale hierarchical image database// 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009. P. 248–255.

23. *Diederik P Kingma, Jimmy Ba. Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980, 2014.*

СВЕДЕНИЯ ОБ АВТОРАХ



ГАЗИЗОВ Марат Рушанович – магистрант, Казанский (Приволжский) федеральный университет, г. Казань.

Marat Rushanovich GAZIZOV – graduate, Kazan (Volga region) Federal University, Kazan.

Email: gazizovmarat@gmail.com



ГРИГОРЯН Карен Альбертович – кандидат экономических наук, доцент, Казанский (Приволжский) федеральный университет, г. Казань.

Karen Albertovich GRIGORIAN – Candidate of Economics, Associate Professor, Kazan (Volga region) Federal University, Kazan.

Email: karigri@yandex.ru

Материал поступил в редакцию 4 июня 2020 года

УДК 004.8

ОПРЕДЕЛЕНИЕ ДЕФЕКТОВ НА СТАЛЬНЫХ ЛИСТАХ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Р. Д. Гаскаров¹, А. М. Бирюков², А. Ф. Никонов³, Д. В. Агнишвили⁴,
Д. А. Хайрисламов⁵

*Институт информационных технологий и интеллектуальных систем,
Казанский (Приволжский) федеральный университет, Казань*

¹ komnatakita@gmail.com, ² herou.public@gmail.com, ³ nikonowalex@gmail.com,
⁴ daniilak@gmail.com, ⁵ danilhayislamov@gmail.com

Аннотация

Сталь в наши дни является одним из важнейших производственных материалов, который используется повсеместно, от медицины до промышленных отраслей. Своевременное обнаружение и распознавание дефектов на стальных листах после проката – одна из ключевых проблем этого производства с учетом его сложности и необходимости затрат большого количества времени на проведение вручную проверок каждого листа и каждой заготовки. Одними из целей настоящей работы были автоматизация и упрощение данного процесса. Для решения соответствующих задач была использована, в первую очередь, модель сверточной нейронной сети под названием UNet, которая уже зарекомендовала себя как отличный инструмент решения таких задач — при высокой результативности она требует меньшего количества учебных данных. В основе этой модели лежат последовательная, производимая в несколько шагов свертка изображения до приемлемого размера (иными словами, сжатие или кодирование), а затем развертка, восстановление изображения к исходному размеру и соотношению сторон, после чего на выходе будет получена маска изображения с классами элементов, которые необходимо было найти. В дополнение к этой нейронной сети в качестве кодирующего (сворачивающего) слоя была использована другая модель — ResNet34, предварительно обученная на датасете (наборе данных) ImageNet1000. В этой модели также был модифицирован выходной слой — вместо 34 слоев с классами на выходе возвращалось лишь 4, что сокра-

тило время обработки и позволило использовать наиболее удачные определения в результатах. Используя данный подход и проведя все необходимые проверки, при подведении итогов, мы получили результат в 94,8% точности определения дефектов на стальных листах.

Ключевые слова: сверточные нейронные сети, нейронные сети, машинное обучение, сталь, искусственный интеллект, UNet, ResNet, определение дефектов, сегментация, классификация.

ВВЕДЕНИЕ

Учитывая высокую конкуренцию в среде производителей стали, фабрикам и производствам просто необходимо выпускать качественный продукт для получения и закрепления хорошей репутации, а также стабильных продаж и обеспечения безопасности конечных потребителей их продукта. Естественно, чтобы можно было гарантировать такое качество продукта, необходимо разработать и внедрить в производство систему контроля качества, иначе говоря, систему мониторинга технологических процессов (далее – СКК). Сама организация этой СКК, говоря конкретно о производстве стали, должна обеспечивать плотное взаимодействие организационных, научных, методологических, технических, информационных и человеко-ресурсных подходов в работе, что не только позволит повысить качество производимой стали, но и уменьшит количество брака, что непосредственным образом благотворно скажется на снижении стоимости производства [1].

Говоря в целом о возможных дефектах и браке стальной продукции, нужно учитывать, что сами свойства материала, процесс проката, а также используемая СКК могут влиять на появление дефектов на листах. В качестве примера можно привести такие повреждения, как царапины, потертости, следы насекомых, сращение листов друг с другом, швы на листах, отслоение частиц с листа, ожоги поверхности листа и прочие [2].

Одна из главных задач СКК — решить, может ли готовая продукция быть выпущенной в продажу, поскольку поставка бракованной партии товара может спровоцировать значительно большие расходы, чем ее утилизация или переработка. Каждый дефект может повлиять на конечное решение о реализации стали

на рынке, например, если на стальном листе будут обнаружены трещины, то данный продукт будет маркирован как брак и не будет допущен к продаже. Если же будет найдено что-то менее существенное, например, небольшие следы шлака или пузырьки воздуха на поверхности листа, то это не будет считаться серьезным недостатком, а сталь пойдет в продажу и дальнейшее производство.

Учитывая все вышесказанное, нами было принято решение разбить работу на две составляющие:

1. Найти и определить дефекты на листах, а также выявить закономерности их появления вкупе с другими дефектами (один дефект может быть причиной появления второго и так далее);
2. Провести эксперимент с использованием нейронных сетей для автоматизации такого рода проверок на качество — смогут ли модель или ансамбль моделей научиться самостоятельно определять дефекты, разбивать их на классы и в дальнейшем, судя по классам, принимать решение о допуске или недопуске стальной продукции к дальнейшей реализации.

1. РАБОТЫ, СВЯЗАННЫЕ С ТЕМОЙ ИССЛЕДОВАНИЯ

Нахождение и определение дефектов на стальных листах с использованием машинного обучения и нейронных сетей далеко продвинулось в последние годы. Некоторые подходы используют модели, которые могут работать в реальном времени, что, однако, требует довольно больших вычислительных мощностей на базе графических процессоров [3], которые позволяют обрабатывать изображения, полученные с видеокамер в большом разрешении. Другой подход, менее требовательный к ресурсам, использует пост-обработку изображений и осуществляется последовательно в три основных этапа:

1. Предварительная обработка изображений;
2. Извлечение классов из объектов на изображении;
3. Классификация полученных артефактов [4].

Один из самых первых методов — RCNN — был исследован и опробован Р. Гиршиком [5], но этот метод не очень хорошо подходил для задач определения объектов на изображении — для его успешной работы требовалась масса времени, поскольку каждое изображение обрабатывалось в исходном разреше-

нии, а каждый найденный объект обрабатывался целиком всей моделью, что сказывалось на производительности не лучшим образом.

Далее, с целью повышения производительности подобных моделей был разработан алгоритм под названием Fast RCNN, который пришел на замену оригинальному RCNN [6]. После него, продолжая наращивать скорость работы, появился Faster RCNN, в основу которого была заложена модель глубокого обучения, работающая с гипотезами о расположении объектов. Эта модель вместо выборочного поиска, использовавшегося в Fast RCNN, строила гипотезы о местонахождении объекта на изображении, а затем извлекала полученные классы и возвращала их.

Учитывая успешность архитектуры RCNN-моделей в задачах определения местонахождения и распознавания объектов на изображениях, нами было принято решение использовать данную архитектуру наряду с UNet, которую также было решено несколько углубить и модифицировать.

3. НАБОР ДАННЫХ

Приступая к практической части работы, в первую очередь необходимо определить, как модели будут взаимодействовать с набором данных (датасетом), то есть какое количество изображений будет достаточным для обучения модели, сколько пикселей потребуется обработать и так далее. Каждое изображение также необходимо разметить — указать в отдельных для каждой фотографии файлах с расширением csv номер дефекта и его маску в формате RLE⁵.

Датасет, который используется для обучения модели, должен также быть подготовлен для большей удобочитаемости компьютером путем замены NaN-значений⁶ в колонке EncodePixels в csv-файле (т. е. генерируем HEX-коды каждого пикселя на изображении для увеличения скорости обработки), изменения наименования файлов и добавления колонки classID, чтобы отслеживать информацию о наличии того или иного дефекта было удобнее.

⁵ https://en.wikipedia.org/wiki/Run-length_encoding

⁶ In computing, NaN, standing for Not a Number, is a member of a numeric data type that can be interpreted as a value that is undefined or unrepresentable, especially in floating-point arithmetic

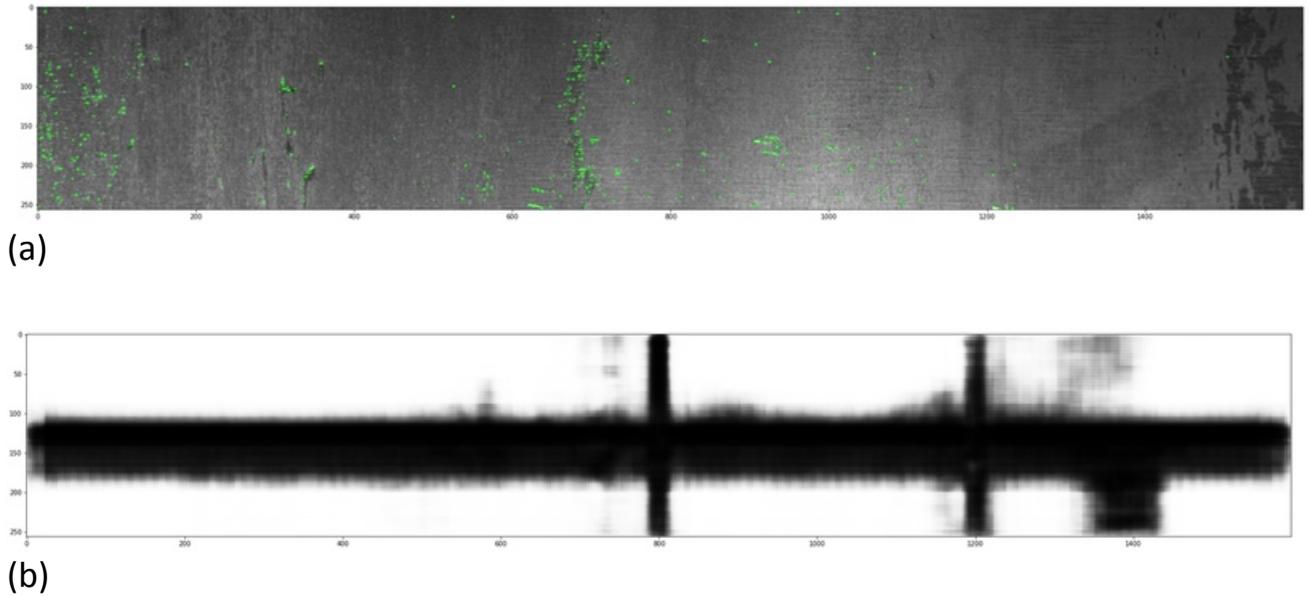


Рисунок 1. Рентгенограмма стального листа с RLE-маской (a) и сама RLE-маска без изображения (b).

Как видно на рис. 1, есть два затруднения в изображениях в датасете:

- на многих изображениях есть довольно большие черные зоны;
- на некоторых изображениях изменена насыщенность цветов.

Однако, учитывая, что на всех фотографиях нет цветов, кроме оттенков черного и серого, их можно проанализировать более простым способом, например, с использованием пиксельной гистограммы.

Как уже упоминалось ранее, если на изображении стального листа есть дефект, то координаты его маски должны быть указаны в строке формата RLE, то есть маска нужного нам дефекта существует с индексацией, наподобие этой: [1234: 1234+3][1239: 1239+16], и так далее. Учитывая это, RLE-строка должна быть декодирована и перенесена на изображения в виде маски⁷.

⁷Kaggle: RLE functions – Run Length Encode & Decode. <https://www.kaggle.com/paulorzp/rle-functions-run-lenght-encode-decode>

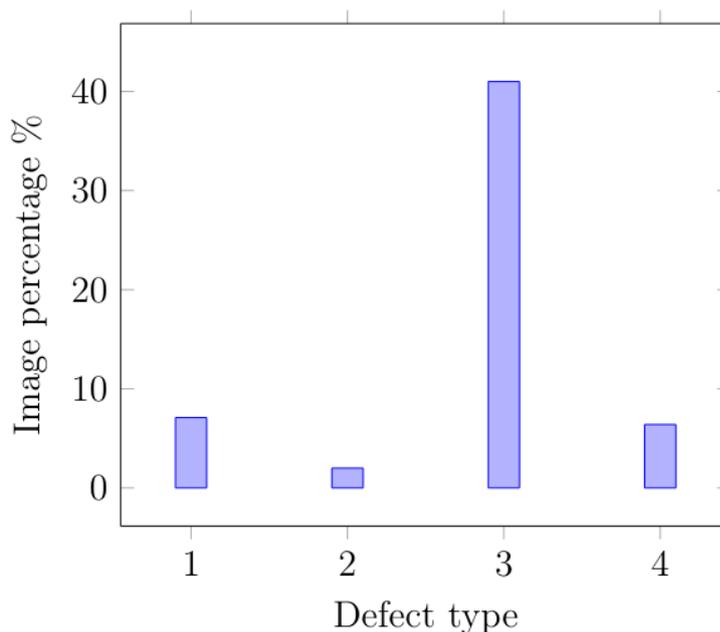


Рисунок 2. Процентаж дефектов по номеру на изображениях в датасете.

Анализ показал, что среди 7095 изображений наибольший процентаж по нахождению имеет дефект №3 (рис. 2). Как видно на рис. 2, распределение дефектов по частоте нахождения неравно. Процент появления дефекта под номером 3 является наивысшим (5150 изображений), в то время как дефект №2 встречается лишь на 2% фотографий во всем наборе данных (247 изображений). Несмотря на это, модель успешно справилась с обучением, поскольку алгоритм работы UNet позволяет точно определять нужные объекты после обучения даже на столь малом наборе изображений.

4. МЕТОДЫ И РЕШЕНИЕ

Основная задача данной работы — избежать возможного человеческого влияния на результат оценки качества стальных листов. Каждый эксперт руководствуется в работе своими собственными принципами и базируется на своем субъективном опыте, поэтому результаты могут отличаться при оценке одного и того же листа разными экспертами. Стоит также учитывать, что люди отличаются от машин и могут сильно нарушить рабочий процесс, если не сорвать его вовсе.

В первую очередь было необходимо автоматизировать процесс обнаружения дефектов. Это было сделано при помощи модели UNet, которая обеспечивает более точную сегментацию при меньшем количестве обучающих данных, что было весьма кстати при очень небольшом количестве изображений в нашем датасете [8].

Модель UNet была представлена в 2015 году Олафом Роннебергером [9] для работ по сегментации биомедицинских изображений и был в дальнейшем модифицирован для использования при поиске дефектов на промышленных производствах. Данная модель использует традиционный подход по свертке, за которым следует так называемое «бутылочное горлышко», а после него — развертка, которая на выходе возвращает маску с сегментированными артефактами. Пропускные соединения между слоями свертки и развертки позволяют восстановить пространственную информацию, которая теряется во время максимального объединения.

Каждый этап развертки состоит из операции увеличения разрешения карты каналов атрибутов, за которой производятся следующие операции:

1. Свертка 2×2 , уменьшающая количество атрибутов каналов;
2. Объединение карты атрибутов каналов, которую необходимо в дальнейшем отделить определенным образом от пути сокращения (это продиктовано потерей крайних пикселей на этапе свертки);
3. Двойная свертка 3×3 , за которой следует активация с выпрямлением (ReLU).

После этих процессов к работе приступает ResNet34 [10] в качестве кодирующей модели. Этот алгоритм достигает 99% точности при использовании на производствах стали горячего проката. ResNet34 обладает также отличной архитектурой для целей извлечения свойств объектов [11]. В модифицированную модель UNet были внедрены блоки ResNet34 на пути свертки, также они действуют по время извлечения пропускных соединений в конце каждого блока на этапе развертки (рис. 3). Входные данные для нейронной сети были подготовлены посредством генерации мозаики 512×512 из входных изображений.

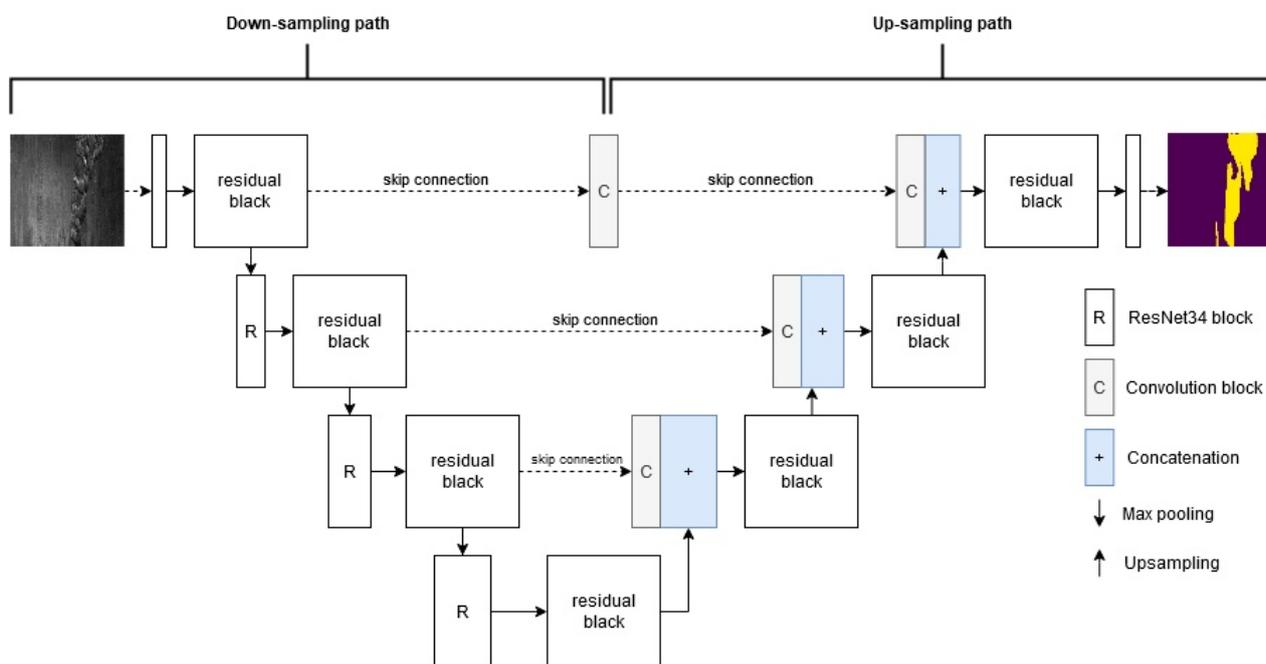


Рисунок 3. Архитектура использованной сверточной нейронной сети. Состоит из двух частей — свёртки и развёртки. Свёртка изображена слева блоками модели ResNet34, а развертка указана справа. Свёртка является стандартным процессом для любой архитектуры сверточных нейронных сетей и состоит из последовательного повторения сжатия 3x3, за которым следуют кодирование длин серий [7] и операция максимального объединения 2x2² для сжатия исходного разрешения изображения.

Dice-коэффициент для тестирования этой сборки выводился после каждой «эпохи» обучения. Когда значение функции потерь оставалось стабильным на одном и том же уровне в течение 10-ти эпох, процесс обучения останавливался. Веса для каждой эпохи обучения сохранялись, и эпоха с лучшим результатом на этапе валидации в дальнейшем использовалась для работы.

На этапе прогнозирования X * X изображений было получено на выходе из модели нейронной сети с определенными отметками надежности прогноза нахождения дефектов (от 0 до 1). Пороговое значение (0,5) было использовано для получения бинарной мозаики предсказания. Мозаичное перекрытие было

использовано для генерации бесшовной сегментации больших зон на изображениях и рентгенограммах [12], а для того, чтобы избежать ненужных артефактов на изображениях и масках, которые часто могут встречаться на краях, использовалась только центральная часть карты предсказаний размером 256x256.

5. ИТОГИ

Основная задача настоящего исследования заключается в создании сверточной нейронной сети, которая сможет на практике подтвердить возможность автоматизации проверок качества стальных листов горячего проката без участия человека в этом процессе. Подводя итог, отметим, эта задача была выполнена в три этапа:

1. Изучение релевантных этой теме нейронных сетей, которые могут потенциально подходить для успешного завершения исследования.
2. Анализ последних научных статей и исследований, а также практических подходов к решению подобных задач с целью определения наиболее эффективных подходов и архитектур нейронных сетей.
3. Разрешение распространенной проблемы с вычислительными мощностями, чтобы сделать возможной обучение и работу модели на доступном оборудовании.

В результате сборки двух моделей — UNet и ResNet34, предобученной на датасете ImageNet1000, были получены следующие результаты:

- 94,8% точность анализа дефектов стальных листов горячего проката.
- Отличные показатели модели на этапе валидации и обучения.

Для оценки полученных результатов была использована метрика Dice-коэффициента. Она демонстрирует сходство или различие между результатом работы нейронной сети и реальными показателями. Чем выше сходство, тем выше коэффициент, также он используется в целях формализации функции потерь в изучении проблем семантической сегментации [13–16]. Как показано на рис. 4, в течение 20 эпох обучения Dice-коэффициент плавно достигает значения около 0,95 и остается на том же уровне.



Рисунок 4. Результаты dice-коэффициента. Эпохи — итерации обучения нейронной сети, Dice score — статистические показатели оценки работы модели, которые определяют, велико ли различие между реальными дефектами на листе стали и теми, которая обнаружила нейронная сеть.

6. ЗАКЛЮЧЕНИЕ

Популярность глубокого обучения в области классификации и обнаружения дефектов с учетом преломления рентгеновских лучей обусловлена высокой результативностью такого подхода [17]. Данный метод был с успехом использован на рентгенограммах и продемонстрировал более высокие результаты, нежели ручная обработка снимков. Учитывая все вышеописанное, внедрение подобных технологий в металлургическую промышленность может существенно снизить расходы на производство, повысить качество продукции и сократить время, требуемое для оценки качества.

В настоящей работе мы предлагаем простой, эффективный и недорогой способ обнаружения и анализа дефектов на листах стали горячего проката, высокая точность работы которого с успехом демонстрирует возможность замены ручного труда при оценке качества продукции на производстве стали.

СПИСОК ЛИТЕРАТУРЫ

1. *Stavropoulos P. et al.* Monitoring and control of manufacturing processes: A review // *Procedia CIRP*. 2013. V. 8. P. 421–425.
2. *Chao W. and al.* Research on Classification of Surface Defects of Hot-rolled Steel Strip Based on Deep Learning // *DEStech Transactions on Computer Science and Engineering*. 2019. 2nd International Conference on Informatics, Control and Automation (ICA 2019). 2019. P. 362–366.
3. *Li J. and al.* Real-time detection of steel strip surface defects based on improved yolo detection network // *IFAC-PapersOnLine*. 2018. V. 51, No. 21. P. 76–81.
4. *Zhang Z., Zhang S., Li Q.* Online Surface Defects Detection System for Cold-rolled Steel Strip // *Recent Patents on Engineering*. 2017. V. 11, No. 1. P. 62–67.
5. *He K. and al.* Mask r-cnn // *Proceedings of the IEEE international conference on computer vision*, 2017. P. 2961–2969.
6. *Ren S. and al.* Faster r-cnn: Towards real-time object detection with region proposal networks // *Advances in Neural Information Processing Systems*, 2015. P. 91–99.
7. *Shang W. and al.* Understanding and improving convolutional neural networks via concatenated rectified linear units // *International Conference on Machine Learning*, 2016. P. 2217–2225.
8. *Ferguson M.K. and al.* Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning // *Smart and Sustainable Manufacturing Systems*. 2018. V. 2. arXiv:1808.02518 [cs.CV]
9. *Ronneberger O., Fischer P., Brox T.* U-net: Convolutional networks for biomedical image segmentation // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2015. P. 234–241.
10. *Wu Z., Shen C., Van Den Hengel A.* Wider or deeper: Revisiting the resnet model for visual recognition // *Pattern Recognition*. 2019. V. 90. P. 119–133.
11. *He K. and al.* Deep residual learning for image recognition // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. P. 770–778.
12. *Ronneberger O., Fischer P., Brox T.* U-net: Convolutional networks for biomedical image segmentation // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2015. P. 234–241.

13. Salehi S.S.M., Erdogmus D., Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks // International Workshop on Machine Learning in Medical Imaging. Springer, Cham, 2017. P. 379–387.

14. Milletari F., Navab N., Ahmadi S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation // 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016. P. 565–571.

15. Wang P., Chung A.C.S. Focal dice loss and image dilation for brain tumor segmentation // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. P. 119–127.

16. Sudre C.H. and al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2017. P. 240–248.

17. Sitdikov V.D. and al. Application of X-ray Technique to Study the Structure of Ultrafine-Grained Ferritic/Martensitic Steel // J. of Materials Engineering and Performance. 2019. V. 28, No. 11. P. 7109–7118.

STEEL DEFECTS ANALYSIS USING CNN (CONVOLUTIONAL NEURAL NETWORKS)

R. D. Gaskarov¹, A. M. Biryukov², A. F. Nikonov³, D. V. Agniashvili⁴,
D. A. Khayrislamov⁵

¹⁻⁵ *Higher Institute of Information Technology and Intelligent Systems*

¹ komnatakita@gmail.com, ² herou.public@gmail.com, ³ nikonowalex@gmail.com,
⁴ daniilak@gmail.com, ⁵ danilhayrislamov@gmail.com

Abstract

Steel is one of the most important bulk materials these days. It is used almost everywhere - from medicine to industry. Detecting this material's defects is one of the most challenging problems for industries worldwide. This process is also manual and time-consuming. Through this study we tried to automate this process. A convolutional neural network model UNet was used for this task for more accurate segmentation with less training image data set for our model. The essence of this NN (neural network) is in step-by-step convolution of every image (encoding) and then stretching them to initial resolution, consequently getting a mask of an image with various classes on it. The foremost modification is changing an input image's size to 128x800 px resolution (original images in dataset are 256x1600 px) because of GPU memory size's limitation. Secondly, we used ResNet34 CNN (convolutional neural network) as encoder, which was pre-trained on ImageNet1000 dataset with modified output layer - it shows 4 layers instead of 34. After running tests of this model, we obtained 92.7% accuracy using images of hot-rolled steel sheets.

Keywords: CNN, neural networks, steel, machine learning, AI, Unet, ResNet, defects detection, segmentation, classification

REFERENCES

1. *Stavropoulos P. et al.* Monitoring and control of manufacturing processes: A review // *Procedia CIRP*. 2013. V. 8. P. 421–425.
2. *Chao W. and al.* Research on Classification of Surface Defects of Hot-rolled Steel Strip Based on Deep Learning // *DEStech Transactions on Computer Science and*

Engineering. 2019. 2nd International Conference on Informatics, Control and Automation (ICA 2019). 2019. P. 362–366.

3. *Li J. and al.* Real-time detection of steel strip surface defects based on improved yolo detection network // IFAC-PapersOnLine. 2018. V. 51, No. 21. P. 76–81.

4. *Zhang Z., Zhang S., Li Q.* Online Surface Defects Detection System for Cold-rolled Steel Strip // Recent Patents on Engineering. 2017. V. 11, No. 1. P. 62–67.

5. *He K. and al.* Mask r-cnn // Proceedings of the IEEE international conference on computer vision, 2017. P. 2961–2969.

6. *Ren S. and al.* Faster r-cnn: Towards real-time object detection with region proposal networks // Advances in Neural Information Processing Systems, 2015. P. 91–99.

7. *Shang W. and al.* Understanding and improving convolutional neural networks via concatenated rectified linear units // International Conference on Machine Learning, 2016. P. 2217–2225.

8. *Ferguson M.K. and al.* Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning // Smart and Sustainable Manufacturing Systems. 2018. V. 2. arXiv:1808.02518 [cs.CV]

9. *Ronneberger O., Fischer P., Brox T.* U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015. P. 234–241.

10. *Wu Z., Shen C., Van Den Hengel A.* Wider or deeper: Revisiting the resnet model for visual recognition // Pattern Recognition. 2019. V. 90. P. 119–133.

11. *He K. and al.* Deep residual learning for image recognition // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. P. 770–778.

12. *Ronneberger O., Fischer P., Brox T.* U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015. P. 234–241.

13. *Salehi S.S.M., Erdogmus D., Gholipour A.* Tversky loss function for image segmentation using 3D fully convolutional deep networks // International Workshop on Machine Learning in Medical Imaging. Springer, Cham, 2017. P. 379–387.

14. *Milletari F., Navab N., Ahmadi S.A.* V-net: Fully convolutional neural networks for volumetric medical image segmentation // 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016. P. 565–571.

15. *Wang P., Chung A.C.S.* Focal dice loss and image dilation for brain tumor segmentation // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. P. 119–127.

16. *Sudre C.H. and al.* Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2017. P. 240–248.

17. *Sitdikov V.D. and al.* Application of X-ray Technique to Study the Structure of Ultrafine-Grained Ferritic/Martensitic Steel // J. of Materials Engineering and Performance. 2019. V. 28, No. 11. P. 7109–7118.

СВЕДЕНИЯ ОБ АВТОРАХ



ГАСКАРОВ Родион Дмитриевич - магистрант, Казанский (Приволжский) Федеральный Университет, г. Казань;

Rodion Dmitrievich GASKAROV — graduate student of Kazan Federal University.

Email: komnatakita@gmail.com



БИРЮКОВ Алексей Михайлович – магистрант, Казанский (Приволжский) Федеральный Университет, г. Казань;

Alexey Mikhailovich BIRYUKOV — graduate student of Kazan Federal University.

Email: herou.public@gmail.com



НИКОНОВ Алексей Федорович – магистрант, Казанский (Приволжский) Федеральный Университет, г. Казань;

Alexey Fedorovich NIKONOV — graduate student of Kazan Federal University.

Email: nikonowalex@gmail.com



АГНИАШВИЛИ Даниил Владиславович – магистрант, Казанский (Приволжский) Федеральный Университет, г. Казань;

Daniil Vladislavovich AGNIASHVILI — graduate student of Kazan Federal University.

Email: daniilak@gmail.com



ХАЙРИСЛАМОВ Данил Айдарович – магистрант, Казанский (Приволжский) Федеральный Университет, г. Казань;

Danil Aydarovich KHAYRISLAMOV — graduate student of Kazan Federal University;

Email: danilhayrislamov@gmail.com

Материал поступил в редакцию 4 августа 2020 года

УДК 004.41 + 004.02

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

А. А. Елизаров¹, Е. В. Разинков²

*Институт вычислительной математики и информационных технологий
Казанского (Приволжского) федерального университета*

¹artelizar@gmail.com, ²evgeny@razinkov.ai

Аннотация

В последнее время активно развивается такое направление машинного обучения, как обучение с подкреплением. Как следствие предпринимаются попытки использования обучения с подкреплением для решения задач компьютерного зрения, в частности для решения задачи классификации изображений. Задачи компьютерного зрения являются на сегодняшний день одними из наиболее актуальных задач искусственного интеллекта.

В статье предложен метод классификации изображений в виде глубокой нейронной сети с использованием обучения с подкреплением. Идея разработанного метода сводится к решению задачи о контекстном многоруком бандите с помощью различных стратегий достижения компромисса между эксплуатацией и исследованием и алгоритмов обучения с подкреплением. Рассмотрены такие стратегии, как ϵ -жадная, ϵ -softmax, ϵ -decay-softmax и метод UCB1, и такие алгоритмы обучения с подкреплением, как DQN, REINFORCE и A2C. Проведен анализ влияния различных параметров на эффективность работы метода.

Ключевые слова: машинное обучение, классификация изображений, обучение с подкреплением, задача о контекстном многоруком бандите.

ВВЕДЕНИЕ

Задача классификации изображений представляет собой одну из важнейших задач компьютерного зрения и заключается в разделении объектов на изображениях на группы, называемые классами [1]. С развитием информационных технологий для решения этой задачи стали успешно применяться методы машинного обучения, в особенности сверточные нейронные сети (англ. convolutional neural network, CNN), которые показывают лучшие результаты по распознаванию объектов на изображениях, начиная с соревнования ILSVRC 2012 (англ. ImageNet Large Scale Visual Recognition Challenge) [2], а в 2015 году превзошли человеческий уровень классификации изображений на наборе данных ImageNet [3].

Обучение глубоких нейронных сетей, в том числе глубоких сверточных сетей, происходит обычно с помощью методов обучения с учителем. Однако в последнее время одним из наиболее развивающихся направлений машинного обучения стало обучение с подкреплением (англ. reinforcement learning). В отличие от обучения с учителем, в обучении с подкреплением модель тренируется, взаимодействуя с окружающей средой и получая от нее отклик [4]. Обучение с подкреплением применяется в различных областях науки и также может использоваться при решении задачи классификации изображений, например, в работах FCAN [5] и DT-RAM [6].

В настоящей работе представлен метод классификации изображений в виде глубокой нейронной сети с использованием обучения с подкреплением, в основу которого также положены интуитивные принципы работы зрительной системы человека. Проведен ряд экспериментов по исследованию влияния различных параметров на эффективность работы метода.

ОПИСАНИЕ РАЗРАБОТАННОГО МЕТОДА

Пусть в качестве базовой модели для решения задачи классификации изображений используется нейронная сеть ResNet [7]. Изображение в виде карт признаков последовательно проходит через остаточные блоки каждой группы ResNet для извлечения от низкоуровневых до высокоуровневых признаков, а далее, на

основе полученных признаков, классифицируется полносвязным слоем. В таблице 1 приведены размеры карт признаков изображений на выходе каждой группы остаточных блоков ResNet.

Таблица 1. Размеры карт признаков изображений на выходе групп блоков ResNet

Номер группы блоков	Размер на выходе
1	56x56x64
2	28x28x128
3	14x14x256
4	7x7x512

Идея разработанного нами метода заключается в использовании помимо сети ResNet дополнительной нейронной сети, которую будем называть агентом. Также будем считать, что сеть ResNet является уже предобученной для используемого набора данных. Задача агента – получив на вход выход группы остаточных блоков ResNet, предсказать, какую область нужно вырезать из полученных карт признаков изображения, чтобы в процессе классификации увеличилась уверенность сети ResNet в принадлежности объекта на исходном изображении правильному классу. Далее карты признаков изображения обрезаются согласно предсказанию агента и подаются на вход следующей группе остаточных блоков ResNet для последующего извлечения признаков и классификации. Такая задача агента ставится из предположения, что при правильном вырезании сеть ResNet сможет лучше классифицировать объект на изображении.

Поставленная задача агента является задачей о контекстном многоруком бандите [4], так как эпизод состоит только из одного действия, и во время обучения агент получает m различных контекстов, где m – количество изображений в обучающей выборке.

Действиями агента являются варианты вырезания части карт признаков изображения. Пусть имеются три возможных типа области вырезания: ограничивающие прямоугольники с размерами сторон, равными соответственно 0.5, 0.75 и 1.0 от размера сторон входных карт признаков, которые являются квадратными. Выбирается параметр шага λ , а затем входные карты признаков замощаются

слева–направо и сверху–вниз каждым из типов прямоугольников с указанным шагом. Замощение ограничивающим прямоугольником с размером сторон 1.0 от размера сторон карт признаков является действием агента «ничего не вырезать».

Вознаграждением агента является разность между уверенностью ResNet в принадлежности объекта на изображении правильному классу, если обрезать карты признаков изображения согласно выбору агента, и уверенностью ResNet в принадлежности объекта на изображении правильному классу, если ничего не вырезать (базовая модель ResNet), или, более формально, вознаграждение $r_{i,t}$ на временном шаге t для i -го изображения из набора данных можно записать как

$$r_{i,t} = outputs_{i,t}(c) - baseline_outputs_{i,t}(c),$$

где $outputs_{i,t}$, $baseline_outputs_{i,t}$ – векторы уверенностей ResNet в принадлежности объекта на i -м изображении различным классам, полученные соответственно после вырезания агентом и при отсутствии вырезания, а c – индекс правильного класса. Таким образом, вознаграждение агента является вещественным числом, изменяющимся в пределах от -1 до 1 .

Агент (нейронная сеть) будет обучаться с помощью выбранного алгоритма обучения с подкреплением, применяя при обучении выбранную стратегию достижения компромисса между эксплуатацией и исследованием (англ. exploitation–exploration) для решения задачи о контекстном многоруком бандите. При обучении агента веса ResNet будут заморожены. Стоит отметить, что для обучения агента в разработанном методе, как и в методах FCAN и DT-RAM, необходимы только метки правильных классов изображений. Место расположения агента в архитектуре ResNet, архитектура агента и другие параметры будут подробно рассмотрены далее.

ЭКСПЕРИМЕНТЫ

Особенности реализации

Для реализации разработанного метода применялся язык программирования Python 3.8 с библиотекой для глубокого обучения PyTorch 1.4 [8]. Для обучения нейронных сетей использовалась среда Google Colaboratory [9].

Описание набора данных

В качестве набора данных для обучения, тестирования нейронных сетей и проведения экспериментов было использовано подмножество набора данных ImageNet [10], который содержит почти 1.4 млн. изображений большого размера, принадлежащих 1000 различным классам от категории животных до категорий мебели и одежды.

Выбор набора данных ImageNet обусловлен наличием большого количества изображений для снижения влияния переобучения и наличием изображений большого размера для проверки эффективности вырезания. Использование только подмножества данных ImageNet обусловлено тем, что весь этот набор данных достаточно большой и для решения задачи не требовался полностью. В подмножество были включены только классы животных, что позволило приблизить исходную задачу классификации к задаче детального распознавания (англ. fine-grained recognition) [11] и увеличить влияние вырезания на эффективность классификации.

Приведем основные характеристики использованного подмножества:

- 60 классов различных животных,
- 14 тыс. изображений для обучения ResNet,
- 55 тыс. изображений для обучения агента,
- 8 тыс. изображений в валидационной выборке.

Основные параметры

Для изображений из использованного набора данных применялись следующие преобразования (указаны соответствующие классы преобразований в библиотеке PyTorch): `transforms.Resize(256)`, `transforms.RandomCrop((224, 224))` при обучении ResNet, `transforms.CenterCrop((224, 224))` при обучении агента.

В качестве базовой модели для классификации изображений использовалась модифицированная нейронная сеть ResNet-50, которая была инициализирована весами, предобученными на ImageNet, и дообучалась на используемом подмножестве. Модификация состояла в использовании перед полносвязным слоем ResNet слоя субдискретизации MaxPool вместо слоя AveragePool. Также в модифицированной архитектуре применялся слой ROI Pool.

Для обеспечения стабильности при обучении агента аугментация данных не применялась. В качестве метода оптимизации использовался метод Adam. Так как скорость обучения в Adam отдельно адаптируется для каждого обучаемого параметра, было принято решение не применять снижение скорости обучения, а использовать Adam с постоянным коэффициентом обучения. В качестве метода регуляризации применялся метод L2 weight decay.

Важным моментом является использование изображений из валидационной выборки агента так же, как и изображений из обучающей выборки агента для обучения ResNet. Этот выбор основан на предположении, что для эффективной работы агента нужно обучать и тестировать его на выходах группы блоков ResNet из одного распределения, то есть, другими словами, и обучать, и тестировать агента необходимо на изображениях, которые либо использовались при обучении ResNet, либо нет (то есть на изображениях, на которые ResNet реагирует схожим образом). Если для обучения и тестирования агента использовать изображения, на которых ResNet не обучалась, то в этом случае может быть сложным исследование влияния гиперпараметров на эффективность работы агента. Поэтому сначала были проведены эксперименты с агентом, который обучался и тестировался на изображениях, использованных при обучении ResNet.

Проверка предположения об эффективности вырезания

Задача агента строится на предположении об увеличении точности классификации за счет вырезания в процессе классификации правильной области на изображении. Для проверки этого предположения применялось моделирование работы «идеального» агента. Принцип работы такого агента заключался в поиске для каждого изображения из набора данных действия, приносящего наибольшее вознаграждение, и последующем использовании этого действия для вырезания в процессе классификации. Такой поиск осуществлялся с помощью метода полного перебора и был возможен благодаря относительно небольшому числу допустимых действий агента. По результатам классификации с использованием «идеального» агента из таблицы 2 можно сделать вывод, что исходное предположение действительно является верным.

Таблица 2. Результаты классификации ResNet с использованием «идеального» агента

Тип вы-борки	Точность клас-сификации без учета вы-резания	Среднее лучшее воз-награжде-ние	Средняя доля действий с поло-жительным воз-награждением	Точность клас-сифика-ции с учетом вырезания
Обучающая	0.8609	0.0865	0.192	0.9539
Валидаци-онная	0.7992	0.0647	0.227	0.8687

Выбор места расположения агента в архитектуре ResNet

В процессе классификации сетью ResNet карты признаков изображений в мини-батче после вырезания агентом могут иметь разный размер, так как у агента есть возможность вырезать области различных размеров. Для вырезания и приведения вырезанных карт признаков к одному размеру применяется слой ROI Pool [12], который часто применяется при решении задачи локализации изображений.

После вырезания карты признаков изображений вероятнее всего будут уменьшены (если только агентом не будет выбрано действие «ничего не вырезать»), поэтому было бы разумным использовать агента и слой ROI Pool между соседними группами остаточных блоков ResNet, так как там, в архитектуре ResNet, происходит уменьшение размера карт признаков с помощью сверточного слоя с шагом свертки, равным 2. При этом шаг свертки при использовании агента в этом месте станет равным 1. На рисунке 1 показана визуализация изображения из набора данных и его карт признаков, которые являются выходами 1-й, 2-й и 3-й групп остаточных блоков ResNet соответственно.

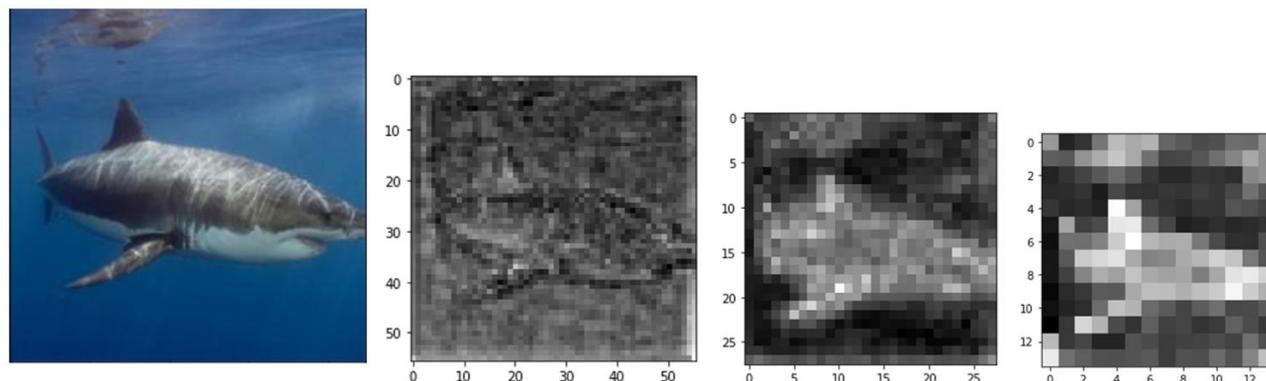


Рисунок 1. Визуализация изображения и выходов 1-й, 2-й и 3-й групп остаточных блоков ResNet.

Из рисунка видно, что после 3-й группы блоков ResNet определяет части объекта, наиболее отличающие его от других, например, у акулы это могут быть плавники. Поэтому использование агента после 3-й группы блоков может позволить в дальнейшем больше сосредоточиться на этих частях. Для упрощения обучения агента будем в дальнейших экспериментах использовать размер шага за-мощения $\lambda=2$, что дает агенту 26 возможных действий.

Архитектура агента

Базовым прототипом для архитектуры агента была архитектура, состоящая из двух полносвязных слоев с нелинейной функцией активации ReLU (таблица 3). Также были разработаны другие версии архитектуры, уменьшающие размерность входа. В таблице 4 приведена архитектура, уменьшающая размер карт признаков изображения с помощью слоя субдискретизации MaxPool, в таблице 5 – архитектура, уменьшающая число каналов карт признаков изображения с помощью сверточного слоя с размером ядра свертки 1x1 и слоем нормализации BatchNorm с функцией активации ReLU, а в таблице 6 – архитектура, уменьшающая как размер, так и число каналов входных карт признаков изображения.

Для архитектуры из таблицы 4 рассматривались варианты уменьшения размера карт признаков изображения до 7x7 (4.1) и 3x3 (4.2), для архитектуры из таблицы 5 – уменьшение числа каналов до 512 (5.1) и 256 (5.2). В архитектуре из таблицы 6 размер карт признаков изображения уменьшается до 7x7, а число каналов – до 512. При этом во всех рассматриваемых архитектурах число нейронов в скрытом полносвязном слое равно 512.

Таблица 3. Полносвязная архитектура

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14x14x1024
1	Полносвязный	ReLU	512	512
2	Полносвязный	–	26	26

Таблица 4. Архитектура с уменьшением размера карт признаков

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14x14x1024
1	Слой субдискретизации (MaxPool)	–	–	7x7x1024 / 3x3x1024
2	Полносвязный	ReLU	512	512
3	Полносвязный	–	26	26

Таблица 5. Архитектура с уменьшением числа каналов карт признаков

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14x14x1024
1	Сверточный	–	1x1, 512 / 1x1, 256	14x14x512 / 14x14x256
2	Слой нормализации (BatchNorm)	ReLU	512 / 256	14x14x512 / 14x14x256
3	Полносвязный	ReLU	512	512
4	Полносвязный	–	26	26

Таблица 6. Архитектура с уменьшением размера и числа каналов карт признаков

Номер слоя	Тип слоя	Функция активации	Размер ядра и число фильтров либо количество нейронов	Размер выхода
0	Вход	–	–	14x14x1024
1	Слой субдискретизации (MaxPool)	–	–	7x7x1024
2	Сверточный	–	1x1, 512	7x7x512
3	Слой нормализации (BatchNorm)	ReLU	512	7x7x512
4	Полносвязный	ReLU	512	512
5	Полносвязный	–	26	26

Графики точности классификации моделей на валидационной выборке в зависимости от используемой архитектуры приведены на рисунке 2. Черной пунктирной линией на рисунке обозначена точность классификации базовой модели ResNet. Так как веса ResNet в процессе обучения агента не обновляются, точность классификации базовой модели не изменяется на каждой эпохе.

По графикам на рисунке 2 можно сделать вывод, что лучшие результаты среди разработанных архитектур показали полносвязная архитектура из таблицы 3 и архитектура с уменьшением числа каналов карт признаков до 512 из таблицы 5 (5.1). Также было исследовано влияние числа нейронов в скрытом полносвязном слое на точность классификации для архитектур из таблиц 3 и 5 (5.1). Были рассмотрены варианты значений этого параметра, равные 256, 512 и 1024 нейронов. Результаты эксперимента приведены на рисунке 3.

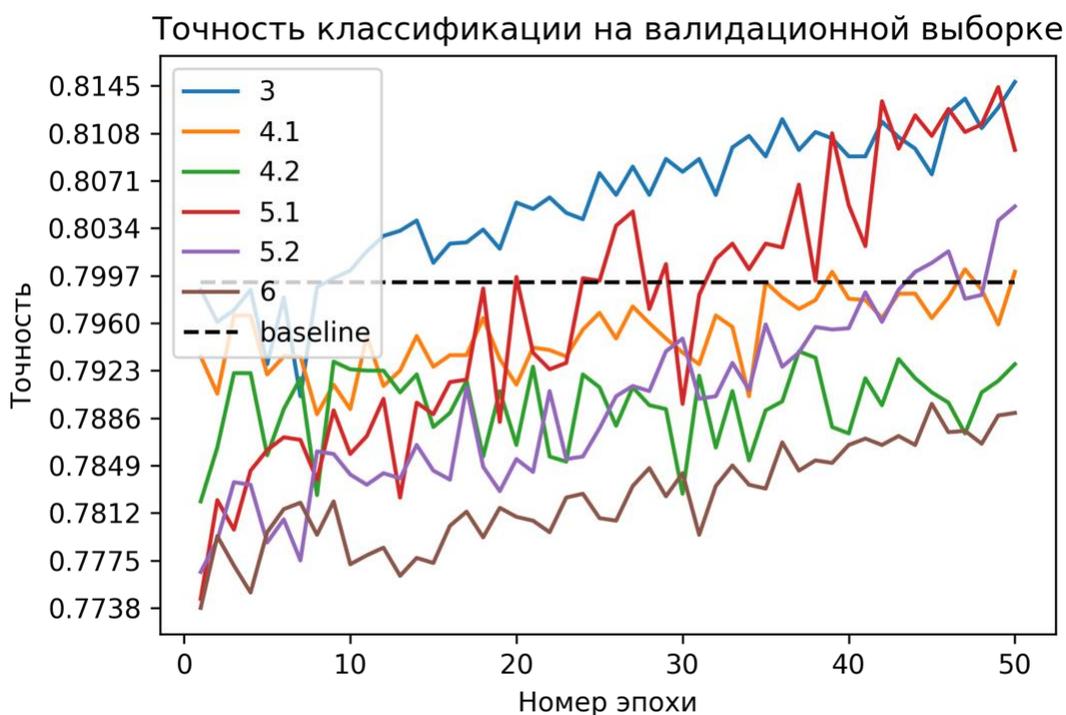


Рисунок 2. Точность классификации моделей в зависимости от используемой архитектуры.

Среди моделей на рисунке 3 наилучший результат показала архитектура 5.1 с 1024 нейронами в скрытом полносвязном слое. Эта архитектура и была выбрана в качестве основной архитектуры агента.

Алгоритм обучения и стратегия выбора действий

В качестве алгоритмов для обучения агента рассматривались алгоритмы обучения с подкреплением DQN [13], REINFORCE [4] и A2C [4]. При этом алгоритм DQN использовался с модификацией воспроизведения опыта [13] с максимально возможным размером буфера, а в функциях потерь алгоритмов REINFORCE и A2C использовался энтропийный бонус [14]. В качестве стратегий выбора действий рассматривались следующие стратегии достижения компромисса между эксплуатацией и исследованием: ϵ -жадная (англ. ϵ -greedy) [4], ϵ -softmax (модификация softmax [4] по аналогии с ϵ -жадной стратегией), ϵ -decay-softmax (по аналогии с ϵ -decay-жадной стратегией [4]) и метод UCB1 [15]. При этом в стратегии ϵ -decay-softmax параметр ϵ уменьшался со временем от значения 1.0 до значения 0.5 в процессе обучения.

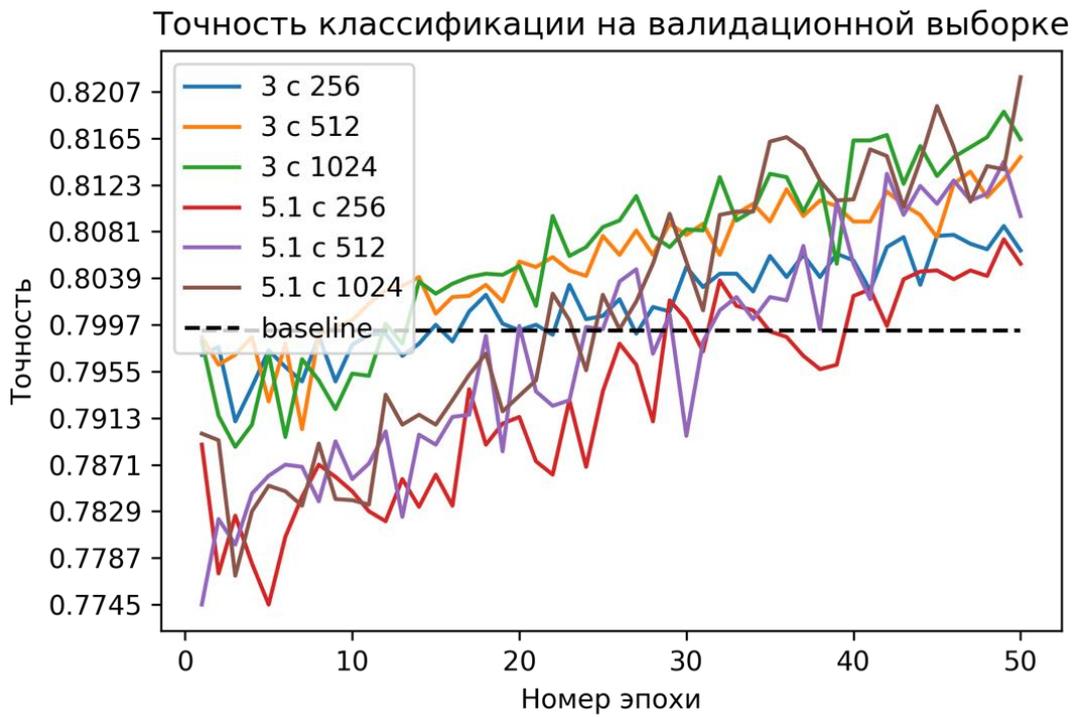


Рисунок 3. Точность классификации лучших моделей в зависимости от числа нейронов в скрытом полносвязном слое.

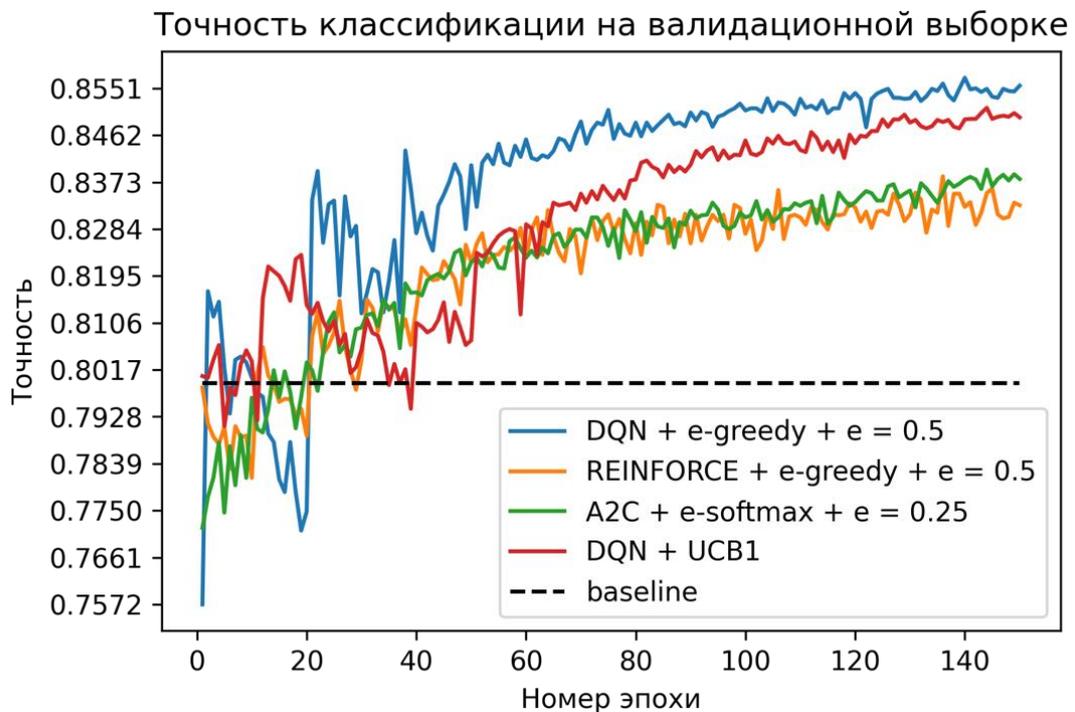


Рисунок 4. Точность классификации лучших моделей.

Для каждого из алгоритмов DQN, REINFORCE и A2C были выбраны наиболее оптимальные стратегии выбора действий, а затем был выбран лучший алгоритм из полученных. Результаты этого эксперимента приведены на рисунке 4 и в таблице 7.

Таблица 7. Результаты классификации лучших моделей

Лучшие модели	Лучшая точность классификации на валидационной выборке
DQN с ϵ -жадной стратегией при $\epsilon = 0.5$	0.8571
DQN с методом UCB1	0.8514
A2C со стратегией ϵ -softmax при $\epsilon = 0.5$	0.8397
REINFORCE с ϵ -жадной стратегией при $\epsilon = 0.5$	0.8392

Таким образом, лучший результат показала модель, обученная с помощью алгоритма DQN и ϵ -жадной стратегии выбора действий с $\epsilon = 0.5$. Она достигла лучшей точности в 0.857 на валидационной выборке, превысив точность базовой модели почти на 0.06, что является почти таким же высоким результатом, как результат классификации «идеального» агента. Результаты работы лучшей модели показаны на рисунках 5 и 6.

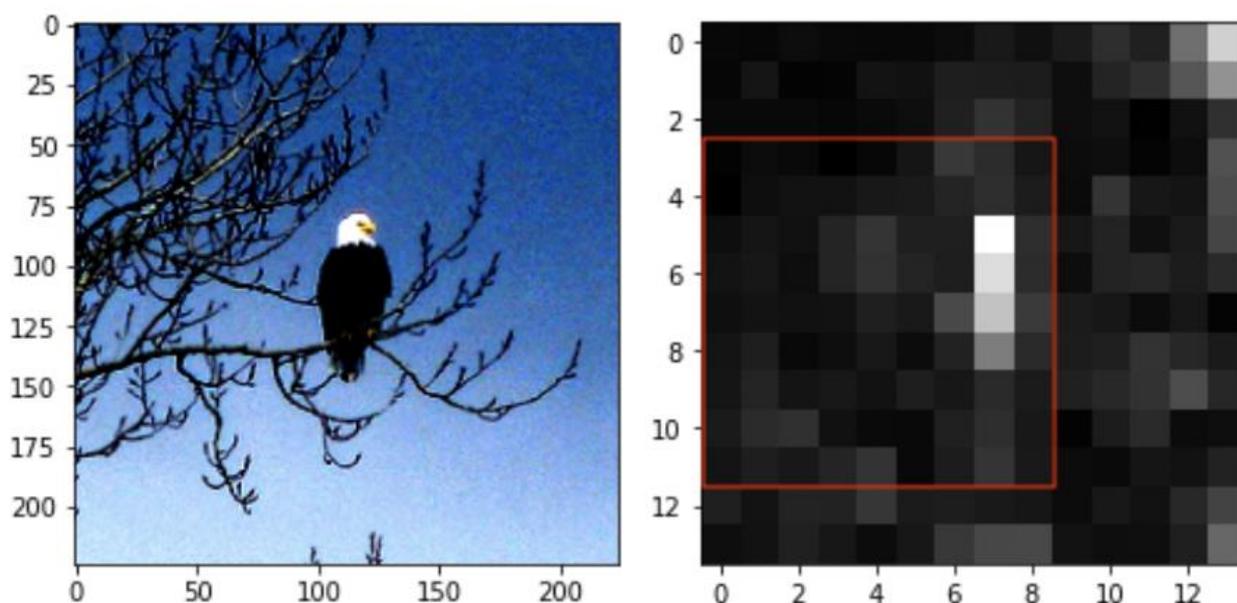


Рисунок 5. Результаты работы лучшей модели на примере.

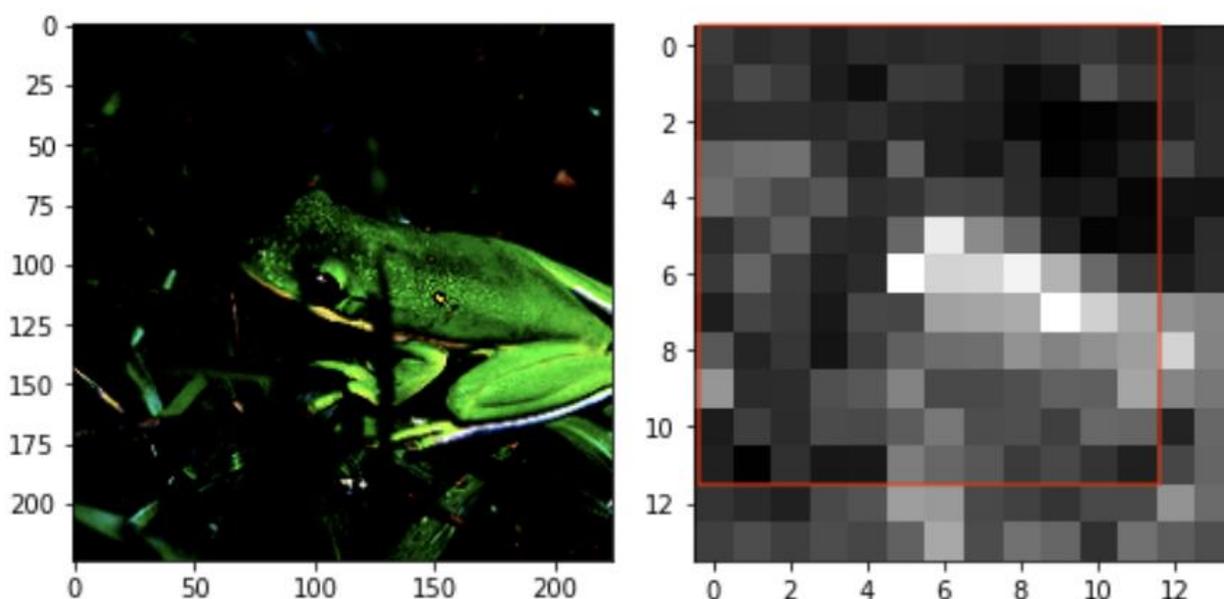


Рисунок 6. Результаты работы лучшей модели на примере.

Использование агента на данных, которые не использовались при обучении ResNet

На практике обучение ResNet на валидационной выборке агента неприменимо. Но можно предложить подход, в котором валидационная выборка используемого набора данных делится на две части: на одной части будет проходить обучение агента, а на другой – его тестирование. При этом сеть ResNet, как и предполагается, не будет обучаться на изображениях из валидационной выборки агента и также на изображениях из обучающей выборки агента. Таким образом, агент будет обучаться и тестироваться на изображениях, на которые ResNet будет реагировать похожим образом, так как при обучении сети ResNet эти изображения использоваться не будут.

В качестве валидационной выборки для этого эксперимента использовалась валидационная выборка полного набора данных ImageNet, которая содержит 50 тыс. изображений 1000 различных классов. Эта выборка была поделена на две части: 45 тыс. изображений для обучения агента и 5 тыс. – для тестирования агента. В качестве параметров агента использовались параметры, полученные в ходе предыдущих экспериментов при использовании агента на изображениях, на которых проходило обучение сети ResNet.

По итогам эксперимента точность классификации базовой модели на тестовой части валидационной выборки составила 0.6934, тогда как точность классификации с учетом вырезания агента составила 0.6942. Можно сделать вывод, что предположение о необходимости использования в качестве входа агента выходов группы блоков ResNet из одного распределения является верным, однако обучение агента на изображениях, которые не используются при обучении ResNet, может потребовать большего объема тренировочных данных, а также детального подбора гиперпараметров.

ЗАКЛЮЧЕНИЕ

В статье предложен метод классификации изображений в виде глубокой нейронной сети с использованием обучения с подкреплением. Основная идея метода строится на предположении об увеличении точности классификации за счет вырезания правильной области на изображении в процессе классификации, что приводит к задаче о контекстном многоруком бандите. Это предположение было успешно подтверждено при помощи моделирования работы «идеального» агента.

Были проведены анализ по выбору места расположения агента в архитектуре ResNet, а также эксперименты по выбору архитектуры агента, алгоритма обучения с подкреплением и стратегии выбора действий при обучении. Помимо основных экспериментов также была осуществлена валидация других гиперпараметров нейронной сети. При проведении экспериментов изображения из валидационной выборки агента использовались при обучении сети ResNet, чтобы оценить эффективность работы агента для входа из выходов группы блоков ResNet из одного распределения. При таком подходе лучшая разработанная модель показала точность классификации, значительно превосходящую точность классификации базовой модели, что дает задел для будущих исследований.

Однако на практике использование изображений из валидационной выборки агента для обучения ResNet неприменимо. Поэтому был предложен подход, в котором изображения из обучающей и валидационной выборок агента не использовались при обучении ResNet. В результате этого эксперимента точность классификации с учетом вырезания агента оказалась немного выше точности базовой модели.

Результаты работы агента сильно отличаются в зависимости от того, на каких данных агент обучается. Это может быть связано со сложностью обучения агента на изображениях, на которых не проходило обучение сети ResNet. Дальнейшая работа по развитию разработанного метода видится в более детальном исследовании эффективности работы агента на изображениях, которые не использовались для обучения сети ResNet.

СПИСОК ЛИТЕРАТУРЫ

1. *Goodfellow I., Bengio Y., Courville A.* Deep learning // С.: The MIT Press, 2016, URL: <https://www.deeplearningbook.org/>.
2. *Krizhevsky A., Sutskever I., Hinton G. E.* ImageNet Classification with Deep Convolutional Neural Networks // Advances in neural information processing systems, 2012. Vol. 25, No. 2. P. 1097–1105, DOI: 10.1145/3065386.
3. *Russakovsky O., Deng J., Su H. et al.* ImageNet Large Scale Visual Recognition Challenge // International Journal of Computer Vision. 2015. Vol. 115, No. 3. P. 211–252, DOI: 10.1007/s11263-015-0816-y.
4. *Sutton R. S., Barto A. G.* Reinforcement learning: An introduction // С.: The MIT Press, 2018. URL: <http://www.incompleteideas.net/book/RLbook2020.pdf/>.
5. *Liu X., Xia T., Wang J. et al.* Fully Convolutional Attention Networks for Fine-Grained Recognition // arXiv:1603.06765, 2017.
6. *Li Z., Yang Y., Liu X. et al.* Dynamic Computational Time for Visual Attention // arXiv:1703.10332, 2017.
7. *He K., Zhang X., Ren S., Sun J.* Deep Residual Learning for Image Recognition // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016. P. 770–778, DOI: 10.1109/CVPR.2016.90.
8. PyTorch, 2016, URL: <https://pytorch.org/>.
9. Google Colaboratory, 2017, URL: <https://colab.research.google.com/>.
10. ImageNet Dataset, 2016, URL: <http://image-net.org/>.
11. Fine-Grained Image Classification, 2019, URL: <https://paperswithcode.com/task/fine-grained-image-classification/>.
12. *Girshick R.* Fast R-CNN // Proceedings of the IEEE International Conference on Computer Vision, 2015. P. 1440–1448, DOI: 10.1109/ICCV.2015.169.

13. *Mnih V., Kavukcuoglu K., Silver D. et al.* Playing Atari with Deep Reinforcement Learning // arXiv:1312.5602, 2013.
 14. *Abdolmaleki A., Springenberg J. T., Degraeve J. et al.* Relative Entropy Regularized Policy Iteration // arXiv:1812.02256, 2018.
 15. *Auer P., Cesa-Bianchi N., Fischer P.* Finite-time Analysis of the Multiarmed Bandit Problem // Machine Learning, 2002. Vol. 47, No. 2-3. P. 235–256, DOI: 10.1023/A:1013689704352.
-

IMAGE CLASSIFICATION USING REINFORCEMENT LEARNING

Artem Elizarov¹, Evgenii Razinkov²

Institute of Computational Mathematics and Information Technology, Kazan Federal University

¹artelizar@gmail.com, ²evgeny@razinkov.ai

Abstract

Recently, such a direction of machine learning as reinforcement learning has been actively developing. As a consequence, attempts are being made to use reinforcement learning for solving computer vision problems, in particular for solving the problem of image classification. The tasks of computer vision are currently one of the most urgent tasks of artificial intelligence.

The article proposes a method for image classification in the form of a deep neural network using reinforcement learning. The idea of the developed method comes down to solving the problem of a contextual multi-armed bandit using various strategies for achieving a compromise between exploitation and research and reinforcement learning algorithms. Strategies such as ϵ -greedy, ϵ -softmax, ϵ -decay-softmax, and the UCB1 method, and reinforcement learning algorithms such as DQN, REINFORCE, and A2C are considered. The analysis of the influence of various parameters on the efficiency of the method is carried out, and options for further development of the method are proposed.

Keywords: machine learning, image classification, reinforcement learning, contextual multi-armed bandit problem.

REFERENCES

1. Goodfellow I., Bengio Y., Courville A. Deep learning // C.: The MIT Press, 2016. URL: <https://www.deeplearningbook.org/>.
2. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks // Advances in neural information processing systems, 2012. Vol. 25, No. 2. P. 1097–1105, DOI: 10.1145/3065386.

3. *Russakovsky O., Deng J., Su H. at all.* ImageNet Large Scale Visual Recognition Challenge // *International Journal of Computer Vision*. 2015. Vol. 115, No. 3. P. 211–252, DOI: 10.1007/s11263-015-0816-y.
4. *Sutton R. S., Barto A. G.* Reinforcement learning: An introduction // C.: The MIT Press, 2018. URL: <http://www.incompleteideas.net/book/RLbook2020.pdf/>.
5. *Liu X., Xia T., Wang J. at all.* Fully Convolutional Attention Networks for Fine-Grained Recognition // arXiv:1603.06765, 2017.
6. *Li Z., Yang Y., Liu X. at all.* Dynamic Computational Time for Visual Attention // arXiv:1703.10332, 2017.
7. *He K., Zhang X., Ren S., Sun J.* Deep Residual Learning for Image Recognition // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016. P. 770–778, DOI: 10.1109/CVPR.2016.90.
8. PyTorch, 2016. URL: <https://pytorch.org/>.
9. Google Colaboratory, 2017. URL: <https://colab.research.google.com/>.
10. ImageNet Dataset, 2016. URL: <http://image-net.org/>.
11. Fine-Grained Image Classification, 2019.
URL: <https://paperswithcode.com/task/fine-grained-image-classification/>.
12. *Girshick R.* Fast R-CNN // *Proceedings of the IEEE International Conference on Computer Vision*, 2015. P. 1440–1448, DOI: 10.1109/ICCV.2015.169.
13. *Mnih V., Kavukcuoglu K., Silver D. at all.* Playing Atari with Deep Reinforcement Learning // arXiv:1312.5602, 2013.
14. *Abdolmaleki A., Springenberg J. T., Degraeve J. at all.* Relative Entropy Regularized Policy Iteration // arXiv:1812.02256, 2018.
15. *Auer P., Cesa-Bianchi N., Fischer P.* Finite-time Analysis of the Multiarmed Bandit Problem // *Machine Learning*, 2002. Vol. 47, No. 2-3. P. 235–256, DOI: 10.1023/A:1013689704352.

СВЕДЕНИЯ ОБ АВТОРАХ



ЕЛИЗАРОВ Артем Александрович – магистрант Института вычислительной математики и информационных технологий Казанского (Приволжского) федерального университета.

Artem Aleksandrovich ELIZAROV – undergraduate of the Institute of Computational Mathematics and Information Technology at Kazan Federal University.

email: artelizar@gmail.com



РАЗИНКОВ Евгений Викторович – к. ф.-м. н., ассистент кафедры системного анализа и информационных технологий Института вычислительной математики и информационных технологий Казанского (Приволжского) федерального университета.

Evgenii Viktorovich RAZINKOV – assistant professor at the Institute of Computational Mathematics and Information Technology at Kazan Federal University, PhD.

email: evgeny@razinkov.ai

Материал поступил в редакцию 12 мая 2020 года

УДК 004

ПЕРСИСТЕНТНЫЕ ГОМОЛОГИИ: ПРИМЕНЕНИЕ К МОНИТОРИНГУ ОБРАЗОВАНИЯ ТРЕЩИН ПРИ ГИДРОРАЗРЫВЕ ПЛАСТА

К. Ю. Ерофеев¹, М. Т. Зиятдинов², Е. В. Мокшин³

Казанский (Приволжский) федеральный университет, Казань

¹krllero@gmail.com, ²gltronred@gmail.com, ³zhen-moks@yandex.ru

Аннотация

Персистентные гомологии – это инструмент топологического анализа данных, отражающий то, как изменяется топологическая структура данных с изменением масштаба их восприятия. В статье дано применение персистентных гомологий к мониторингу образования трещин при гидроразрыве пласта, что позволяет исследователю естественным образом учитывать априорную информацию.

Ключевые слова: топологический анализ данных, персистентные гомологии, гидроразрыв пласта, микросейсмические события.

ВВЕДЕНИЕ

Теория персистентных гомологий начала развиваться в конце прошлого тысячелетия независимо тремя группами ученых: Патрицио Фроссини и Массимо Ферри из итальянской Болоньи, Ванессой Робинс из Колорадо в ее докторской диссертации и в биогеометрических исследованиях Герберта Эдельсбурнера из Северной Каролины [1]. Группа, в которую входили Патрицио Фроссини и Массимо Ферри, изучала нульмерные гомологии как теорию размера, эти исследования были мотивированы изучением естественных псевдо-расстояний между двумя функциями на гомеоморфных топологических пространствах.

Идея альфа-формы, предложенная в 1983 году [2], лежала в основе докторской диссертации Ванессы Робинс, применившей ее к изучению фракталов [3]. В 2002 году был разработан и представлен алгоритм вычисления персистентных гомологий [4].

Давид Кохен-Штайнер, Герберт Эдельсбурнер и Джон Харер в 2007 году [5] показали стабильность персистентных гомологий в том смысле, что небольшие

изменения в данных ведут к небольшим изменениям баркодов соответствующих гомологий.

Иногда, чтобы отобразить представляющие интерес признаки в элементы гомологий, в существующее пространство данных добавляется дополнительное построение. Так, в [6] Ксяосинь Ю рассматривает предложения текста как точки в евклидовом пространстве, основываясь на модели «мешок слов», и обращает внимание на то, что близкие по смыслу предложения находятся в данном представлении ближе друг к другу. Ксяосинь Ю поставил задачу получить с помощью первых гомологий информацию о количестве отсылок в тексте. Чтобы отобразить отсылки в тексте в элементы первых гомологий, он достраивает комплекс, соединяя соседние по расположению в тексте предложения 1-симплексами. Важно заметить, что такая процедура не только отображает в элементы первых гомологий все отсылки в тексте, но и делает элементы первых гомологий, соответствующие отсылкам в тексте, тем устойчивее, чем дальше расположены друг от друга отсылка и источник. Описанный метод был протестирован Ксяосинем Ю на эссе, что были написаны детьми и подростками. Предположение, что эссе подростков имеет более сложный смысловой узор, значит, и большее количество отсылок, подтвердилось полностью, так что устойчивость элементов первых гомологий фильтраций, построенных по текстам подростков, значительно превосходила устойчивость элементов первых гомологий в текстах детей.

В настоящее время персистентные гомологии как инструмент топологического анализа данных применяются к широкому спектру задач, среди которых: изучение зрительной коры головного мозга [7], беспроводных сетей [8], сегментирования изображений [9], космологий [10], предвыборных махинаций [11].

В настоящей статье описано применение персистентных гомологий к задаче локализации микросейсмических событий. В первом разделе изложены цели и проблемы микросейсмического мониторинга, во втором разделе представлены основные понятия теории персистентных гомологий и их интерпретация. В третьем разделе описаны применение персистентных гомологий к задачам микросейсмического мониторинга, важную часть которого составляет введение метрики на пространстве микросейсмических событий, которая обсуждается в разделе 4. Результаты соответствующих вычислительных экспериментов приведены

в разделе 5.

1. МИКРОСЕЙСМИЧЕСКИЙ МОНИТОРИНГ

В последнее десятилетие наблюдается увеличение интереса к микросейсмическому мониторингу со стороны нефтегазовой промышленности.

В процессе разработки месторождения геологическая среда подвергается различного рода воздействиям, которые изменяют напряженно-деформированное состояние породы. Одной из наиболее важных подобных процедур является гидравлический разрыв пласта (ГРП).

ГРП – это один из методов повышения продуктивности нефтяного или газового месторождения. Такая процедура заключается в нагнетании жидкости и увеличении давления в скважине на необходимой глубине резервуара с целью создания высоко-проводимой системы трещин, по которым осуществляется приток углеводорода к стволу скважины.

Процедура разрыва пласта является технологически сложной операцией. В процессе ее проведения нужно понимать и контролировать в том числе реакцию пласта на закачку флюида и размещение пропанта внутри образовавшихся трещин. Результаты проведения этой процедуры являются опорой для планирования дальнейших шагов по эксплуатации месторождения.

Разрушения или скольжение породы, происходящие внутри пласта, являются источником акустической эмиссии, представляющей собой микросейсмические события. Микросейсмическая активность содержит информацию о происходящих сдвигах, раскрытиях и закрытиях трещин.

Наиболее эффективной технологией контроля качества ГРП является микросейсмический мониторинг [12–15]. Он осуществляется с помощью сейсмоприёмников, расположенных в соседней скважине или на дневной поверхности, и сопровождает всю процедуру, записывая приходящие колебания. Регистрация микросейсмических возмущений, возникающих при образовании и росте систем трещин, позволяет определить их свойства и рассчитать место и время их зарождения [16].

Микросейсмическое событие характеризуется тензором сейсмического момента (ТСМ) [17]. ТСМ описывается системой из девяти пар сил и отражает деформацию среды, происходящую в месте очага возмущения (рис. 1).

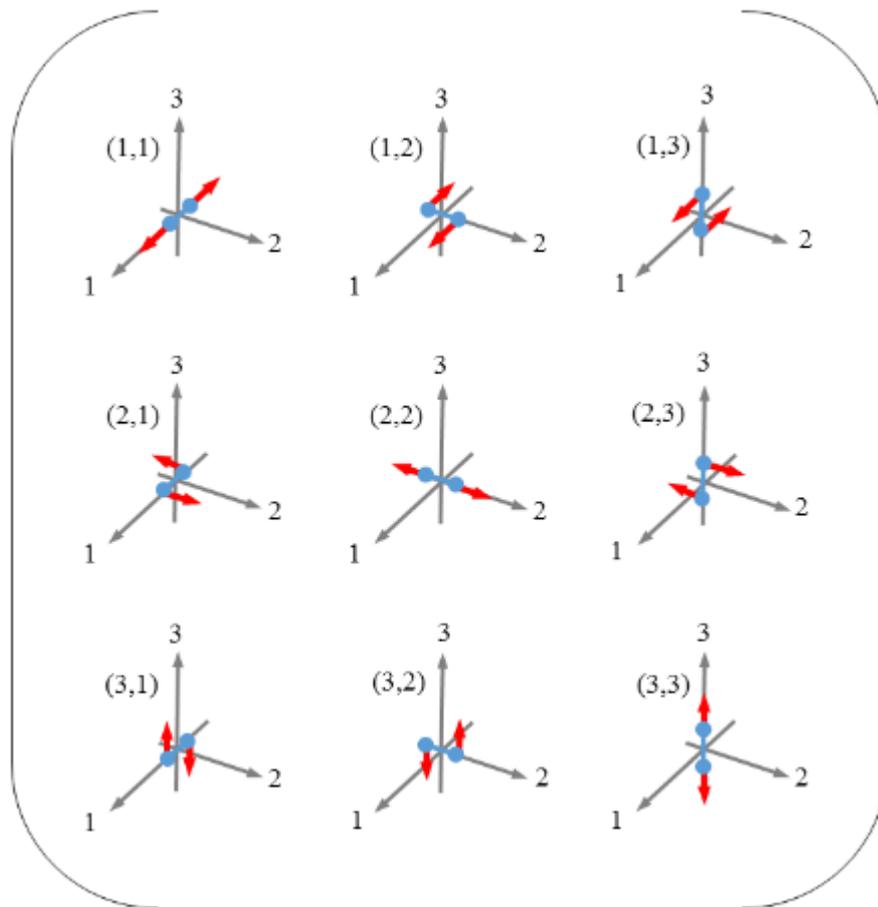


Рисунок 1. Тензор микросейсмического момента.

Тензор сейсмического момента является симметричной и вещественной матрицей, которую всегда можно привести к диагональному виду. Таким образом, конкретное микросейсмическое событие описывается тремя собственными значениями и собственными векторами. Собственное значение отражает амплитуду события по направлению соответствующего ему собственного вектора.

Результатом микросейсмического мониторинга является микросейсмическая активность, накопленная за время проведения ГРП. Однако далеко не все восстановленные события имеют прямое отношение к росту трещин.

Изменение давления в пласте приводит к линейной диффузии порового давления и вызывает микросейсмические события, которые непосредственно не связаны с образованием трещин [13]. Кроме того, сейсмические приемники, кроме полезной информации, регистрируют также проходящие шумы. Шумы имеют разную причину возникновения, в том числе исходят от работающего на

поверхности оборудования, осуществляющего процесс ГРП. Наличие шумов вносит искажения в получаемый результат.

Таким образом, интерпретация полученных событий является важной и сложной процедурой, требующей учета различных факторов. На инженера, который проводит интерпретацию, ложится ответственность за правильную классификацию и отбраковку ложных событий. Его задача – увидеть в сложной картине из множества событий, каждое из которых описывается шестью компонентным тензором, взаимосвязанные физические процессы.

Среди трудностей, с которыми сталкивается инженер, оценивая картину микросейсмических событий, следует отметить следующие.

Многомерность данных: для каждого из микросейсмических событий известны 3 его пространственные координаты, 3 собственных значения, определяющие тип микросейсмического события, и как минимум один главный собственный вектор, определяющий направление.

Направление развития трещины при микросейсмических событиях разного типа имеет разную зависимость от главного собственного вектора тензора микросейсмического момента.

Типы микросейсмических событий не определяются точно, вместо этого мы имеем события, которые с разной степенью уверенности можно отнести к нескольким типам. Таким образом, исследователю необходимо с разной вероятностью учитывать принадлежность события к разным типам и интерпретировать это как свидетельство развития трещины в разных направлениях.

Часто необходимо учитывать *априорную информацию* о направлениях развития трещин или особенности метода микросейсмического мониторинга. Некачественное выполнение данного этапа сильно обесценивает вложенные затраты и труд, поэтому остро встает вопрос о необходимости создания автоматического инструмента, который присваивал бы двум областям микросейсмических событий числовую характеристику, которая может быть интерпретирована как правдоподобность существования между ними трещины, и одновременно с этим допускал бы естественный учет априорной информации.

2. ПЕРСИСТЕНТНЫЕ ГОМОЛОГИИ

Персистентные гомологии передают информацию об узоре пустот в дан-

ных путем идентификации разнообразия классов эквивалентности замкнутых путей в данных, рассмотренных с разным масштабом. К одному классу эквивалентности принадлежат замкнутые пути, которые могут быть переведены друг в друга непрерывным отображением. Такие замкнутые пути называют гомотопными. Ниже приведены основные определения, взятые из [6].

Определение 1. Симплициальный комплекс K – это конечное множество симплексов, такое, что $\sigma \in K, \tau \subset \sigma \Rightarrow \tau \in K$, и если $\sigma, \tau \in K$, то $\sigma \cap \tau$ или пусто, или является гранью σ и τ одновременно.

Определение 2. k -ая группа цепей C_k комплекса K – это свободная абелева группа на множестве ориентированных k -симплексов, принадлежащих этому комплексу, где выполняется $[\sigma] = -[\tau]$, если $\sigma = \tau$ и при этом $[\sigma]$ и $[\tau]$ противоположно ориентированы. Ее элементы – k -цепи, $c \in C_k, c = \sum_i n_i [\sigma_i]$, где $\sigma_i \in K, n_i \in \mathbb{Z}$.

Определение 3. Оператор взятия границы – это гомоморфизм групп $\partial_k: C_k \rightarrow C_{k-1}$, действие которого на элементы C_k определяется линейно через его действие на k -симплексы $\sigma_k = [v_0, v_1, \dots, v_k]$:

$$\partial_k \sigma = \sum_i (-1)^i [v_0, v_1, \dots, \hat{v}_i, \dots, v_k].$$

Через \hat{v}_i обозначена вершина, которая исключается из симплекса.

Определение 4. Операторы взятия границ соединяют группы цепей в цепные комплексы

$$C_*: \dots \xrightarrow{\partial_{k+1}} C_{k+1} \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_{k-1}} \dots .$$

Определение 5. С помощью оператора взятия границы определяются также две важные подгруппы k -х групп цепей: группа циклов $Z_k = \ker \partial_k$ и группа границ $B_k = \text{im } \partial_{k+1}$.

Определение 6. k -ая группа гомологий – это фактор k -ой группы цепей по k -ой группе границ: $H_k = Z_k/B_k$.

Такому определению можно дать следующую интерпретацию: Z_k – это множество всех k -цепей, которые оператор взятия границы переводит в нуль, т. е. множество всех замкнутых контуров, а B_k – это множество всех тех k -цепей, которые являются границами $k+1$ -цепей. Тогда k -ая группа гомологий $H_k =$

Z_k/B_k – это группа всех замкнутых контуров, факторизованная по тем замкнутым контурам, которые являются границами некоторой $k + 1$ -цепи, т. е. окружают некоторый существующий в $k + 1$ -ом измерении кусок пространства, значит, не образующий пустоту. Эта интерпретация объясняет способность гомологий отображать структуру пустот топологического пространства.

Определение 7. Фильтрацией комплекса K называется последовательность вложенных друг в друга комплексов $\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_m = K$. Для общности мы считаем $K_i = K_m$ для всех $i \geq m$. K в таком случае называется фильтрованным комплексом.

Теперь для преобразования наших данных в фильтрацию мы можем интерпретировать данные как точки в пространстве \mathbb{R}^n , а затем можем построить комплекс Чеха или комплекс Рипса с целью захватить топологические инварианты пространства.

Определение 8. Комплекс Чеха $\check{C}_\varepsilon(X)$ для $\varepsilon > 0$ и облака точек X определяется следующим образом: точки из X образуют множество точек комплекса, т. е. множество нульмерных симплексов, а подмножество $\sigma \subset X$ принадлежит комплексу $\sigma \in \check{C}_\varepsilon(X)$, если множество шаров с центром в точках из σ и радиусом ε имеют непустое пересечение.

Определение 9. Комплекс Рипса $R_\varepsilon(X)$ для $\varepsilon > 0$ и облака точек X строится следующим образом: некоторое подмножество вершин $\sigma \subset X = \{x_0, x_1, \dots, x_k\}$ составляет симплекс в $R_\varepsilon(X)$, если все $x \in \sigma$ попарно близки, т. е.

$$d(x_i, x_j) < \varepsilon, \forall x_i, x_j \in \sigma.$$

Построение комплекса Чеха – вычислительно более дорогая операция, нежели построение комплекса Рипса, однако комплекс Чеха всегда гомотопически эквивалентен топологическому пространству, полученному как объединение шаров, в то время как для комплекса Рипса это не всегда так [14].

Как для комплекса Чеха, так и для комплекса Рипса мы можем сопоставить набору точек в пространстве естественную последовательность вложенных друг в друга комплексов, так как в обоих случаях выполняется $\varepsilon < \gamma \Rightarrow R_\varepsilon(X) \subseteq R_\gamma(X)$, следовательно, любая возрастающая последовательность неотрицательных ε даст нам фильтрацию. Теперь элемент гомологий, появившийся на эле-

менте фильтрации i и закончивший свое существование на элементе фильтрации p , определяется следующим образом:

$$H_k^{i,p} = Z_k^i / (B_k^{i+p} \cap Z_k^i).$$

Графически гомологии изображаются с помощью специальных графиков, называемых баркодами. По оси Y на баркодах отображается порядковый номер элемента гомологий, а координаты начала и конца баркода по оси X соответствуют моменту появления i и исчезновения $i + p$ соответствующего элемента гомологий.

Так, например, для набора точек, изображенного на рис. 2, с фильтрацией, построенной с помощью комплекса Рипса и евклидовой метрики, баркод, отображающий значимые элементы первых гомологий, будет выглядеть так, как показано на рис. 3. Здесь наибольший по ширине баркод соответствует вкладу в первые гомологии большого круга, второй по ширине – вкладу маленького, оставшиеся баркоды с маленькой продолжительностью жизни соответствуют шуму.

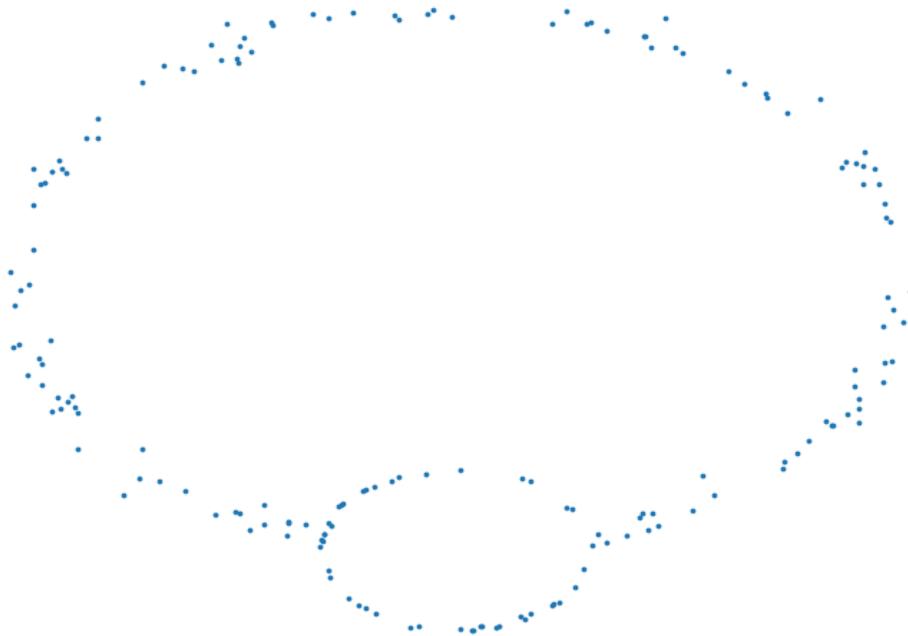


Рисунок 2. Точки, образующие два круга: простейший пример для иллюстрации работы персистентных гомологий.



Рисунок 3. Баркод персистентных гомологий для данных, изображенных на рис. 2.

3. ПРИМЕНЕНИЕ К МОНИТОРИНГУ ТРЕЩИН ГРП

Одной из главных трудностей в ручной обработке данных микросейсмического мониторинга является их многомерность. Так, микросейсмические события как правило имеют три пространственных измерения и шесть измерений, составленных из координат главного собственного вектора и собственных чисел тензора микросейсмического момента. В данном разделе для более удобной и наглядной визуализации данных будут рассмотрены микросейсмические события на плоскости, однако представленные результаты естественным образом переносятся на случай большего количества размерностей.

Чтобы отыскать трещины в наборе микросейсмических событий, мы выделяем две области, наличие трещины между которыми будем исследовать. По окончании работы представленного алгоритма мы получим число, которое можно будет интерпретировать как правдоподобность наличия трещины между выделенными областями.

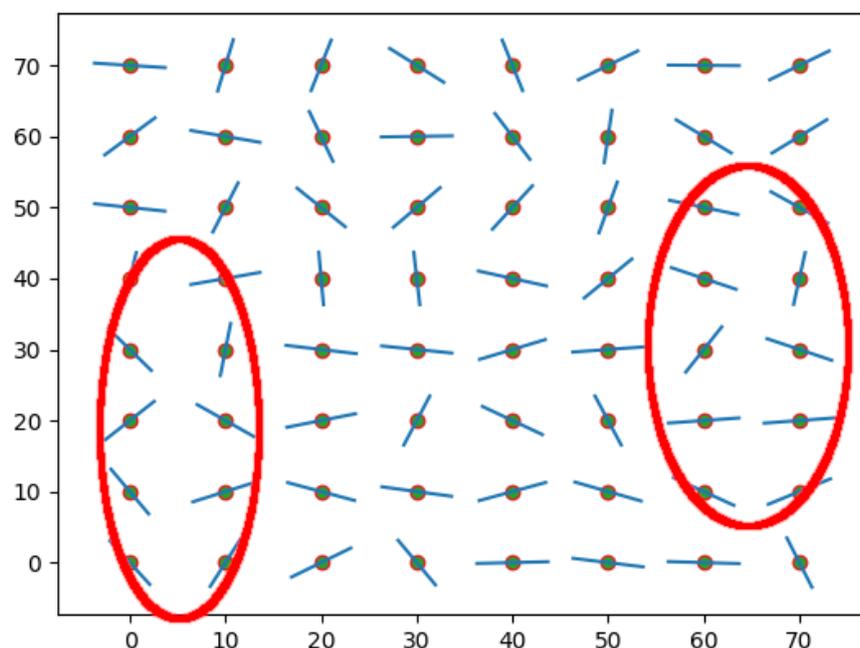


Рисунок 4. Микросейсмические события с двумя выделенными областями интереса.

Например, на графике (рис. 4) положение отрезка в пространстве характеризуют положение тензора микросейсмического момента в пространстве, а его направление – направление наибольшего собственного вектора тензора. Красными эллипсами выделена область данных, между которыми будет производиться поиск трещины. Тип микросейсмического события на графике не отображается.

Чтобы построить по набору микросейсмических событий фильтрацию, мы вводим на множестве микросейсмических событий метрику, учитывающую не только пространственную близость микросейсмических событий друг к другу, но и их тип и направление их собственных векторов так, чтобы близкие в этой метрике события наиболее вероятно составляли участок трещины.

Также мы добавляем к нашим данным точку, которая будет в выбранной нами метрике располагаться близко к двум выделенным множествам точек, после чего трещины, которые проявляются как линии из микросейсмических событий, близко расположенных в метрике, учитывающей собственные вектора тензора микросейсмического момента, будут проявляться как элементы первых го-

мологий. А наибольшая продолжительность жизни элемента первых гомологий может быть интерпретирована как правдоподобность наличия трещины между двумя выделенными областями данных.

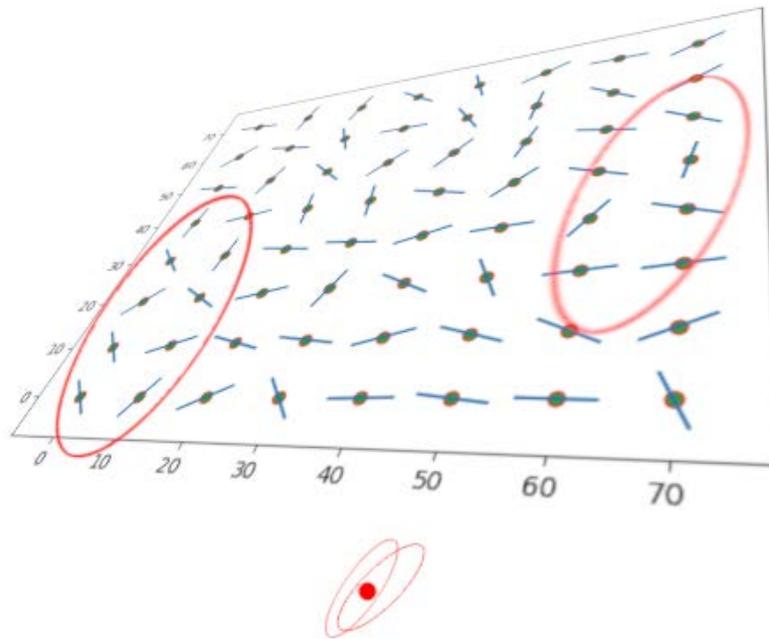


Рисунок 5. Построение, обеспечивающее отображение трещин в элементы гомологий.

На графике (рис. 5) визуализировано добавление дополнительной точки, которая обеспечит отображение трещин, соединяющих две конкретные выделенные области в элементы первых гомологий.

Для остальных точек пространства мы строим нашу метрику как сумму метрик: евклидовой метрики на пространстве физического расположения микросейсмических событий и некоторой метрики на пространстве тензоров микросейсмического момента, умноженной на некоторый коэффициент, что без ограничения общности может быть записано как

$$d_M(M_x, M_y) = d_L(L_x, L_y) + h d_E(E_x, E_y),$$

где d_M – метрика на пространстве микросейсмических событий, d_L – евклидова метрика на координатах микросейсмических событий, d_E – метрика на тензорах микросейсмического события.

4. МЕТРИКА НА ТЕНЗОРАХ МИКРОСЕЙСМИЧЕСКОГО СОБЫТИЯ

В первую очередь отметим, что для успешного построения фильтрации используемая функция расстояния необязательно должна быть метрикой в алгебраическом смысле, а именно, необязательным является выполнения правила треугольника, в то время как необходимыми являются положительность и коммутативность метрики.

В процессе построения метрики мы учтем как направление трещины, так и ту степень, с которой то или иное событие свидетельствует об определенном направлении трещины.

При мониторинге ГРП исследователя, как правило, интересуют только два вида событий: Tensile Crack и Double-Couple. Double-Couple характеризует собой сдвиг, такие события возникают первыми, когда нагнетаемое давление еще недостаточно, чтобы преодолеть сдавливающие напряжения. Когда давление достигает значения напряжения, возникает трещина отрыва, такие события относят к типу Tensile Crack. На финальном этапе ГРП давление стравливают, что приводит к закрытию некоторых трещин, эти события тоже описываются Tensile Crack, однако направление главного момента будет с противоположным знаком.

Наличие двух видов интересующих нас событий дает два вида зависимости направления развития трещины от наибольшего собственного вектора тензора микросейсмического момента, что отображено на иллюстрациях ниже.

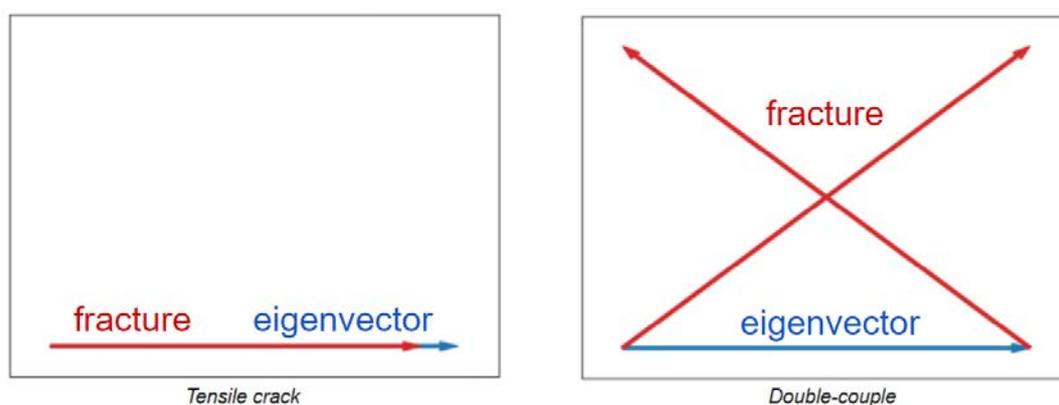


Рисунок 6. Микросейсмические события.

Также мы не можем с полной уверенностью отнести событие к конкрет-

ному типу, на практике мы чаще понимаем, что конкретное событие может быть интерпретировано с разной степенью уверенности как событие разных типов.

Чтобы учесть все возможные интерпретации, расстояние между двумя тензорами микросейсмического момента определяется как минимальное расстояние по всем интересующим нас интерпретациям:

$$d_E(M_x, M_y) = \min d_t(M_x, M_y), t \in [DC, TC - DC, DC - TC, TC]. \quad (1)$$

Через d_{DC} мы обозначили расстояние, рассчитанное, исходя из интерпретации обоих событий как Double-Couple, d_{TC-DC} – исходя из интерпретации события M_x как Tensile Crack и события M_y – как Double-Couple и т. д. Рассмотрим вопрос о присвоении значения параметру h из (1). Ниже представлен способ эвристической оценки значения параметра h . Рассмотрим следующий набор микросейсмических событий:

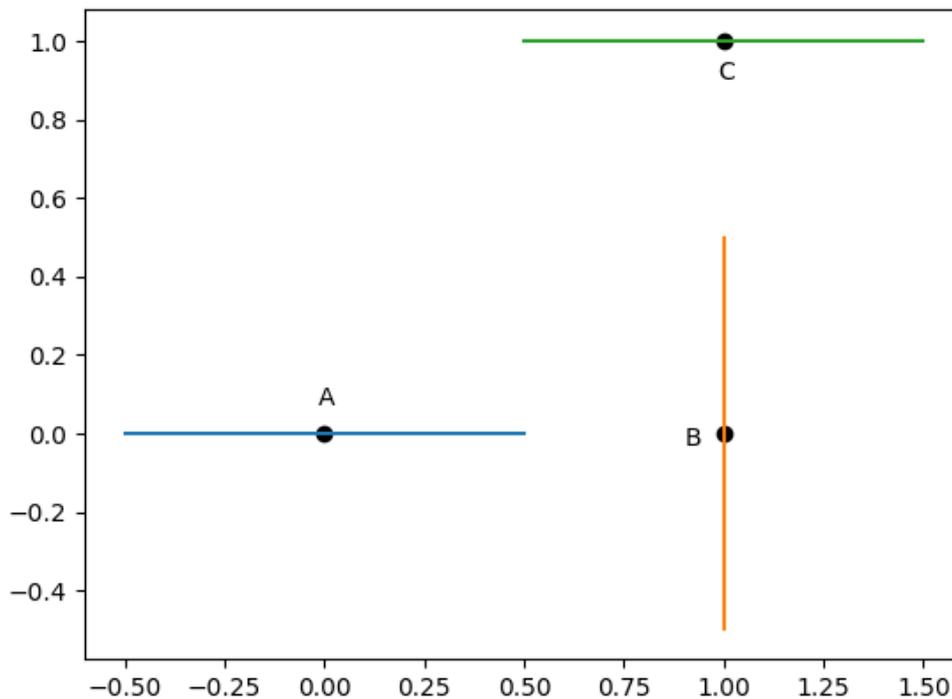


Рисунок 7. Иллюстрация значения метрики.

Здесь микросейсмические события расположены в вершинах квадрата, направление развития трещины, предсказанное событием В, перпендикулярно

направлениям, предсказанным событиями А и С. Исходя из опыта микросейсмических исследований, можно заключить, что $d(A, B) > d(A, C)$, таким образом мы получаем

$$d_E(A, B) = \sqrt{2},$$

$$d_L(A, C) - d_L(A, B) = \sqrt{2} - 1,$$

из чего следует оценка снизу:

$$h \geq d_L(A, B) \frac{\sqrt{2}-1}{\sqrt{2}},$$

где $d_L(A, B)$ – длина шага сетки.

Важно отметить, что манипуляции с метрикой позволяют нам вводить априорную информацию: так, если исследователь не ожидает в своем наборе микросейсмических событий типа DC, то возможность интерпретировать события как событие типа DC можно исключить из описанной выше метрики. Также метрика может быть сжата вдоль оси, по которой наличие трещины более вероятно.

5. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

Для простейшего подтверждения работоспособности предложенного алгоритма были сгенерированы четыре набора событий типа Tensile Crack, с разной степенью очевидности наличия трещины между выделенными областями. Рассмотрим, как ведут себя элементы первых гомологий.

Идеальная трещина отображается в длинный выступающий элемент первых гомологий (рис. 8).

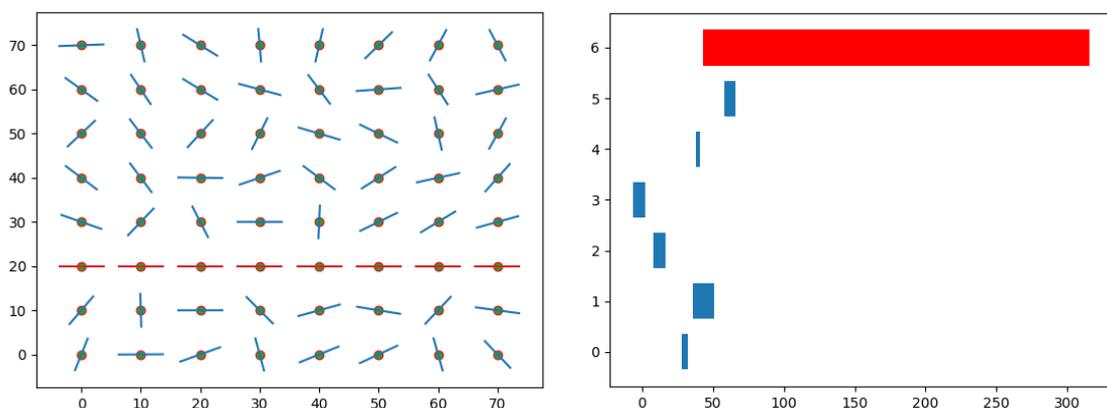


Рисунок 8. Эксперимент без шума.

Трещина с нормальным шумом, добавленным в направления собственных векторов, с дисперсией в 10% от перпендикулярного направления вектора примерно в два раза уменьшает длину жизни элемента первых гомологий (рис. 9).

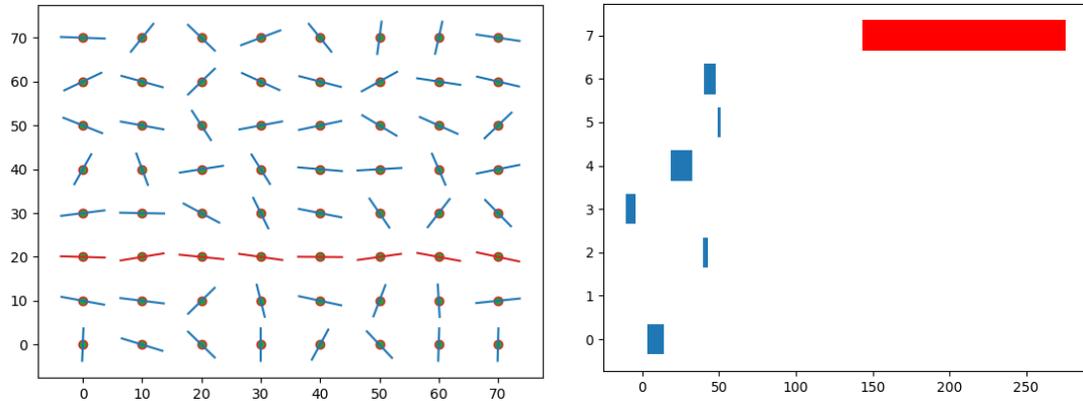


Рисунок 9. Эксперимент с низким уровнем шума.

Трещина с нормальным шумом с дисперсией в 60% еще в два раза уменьшает длину жизни наибольшего элемента первых гомологий, хотя и оставляет его различимым (рис. 10).

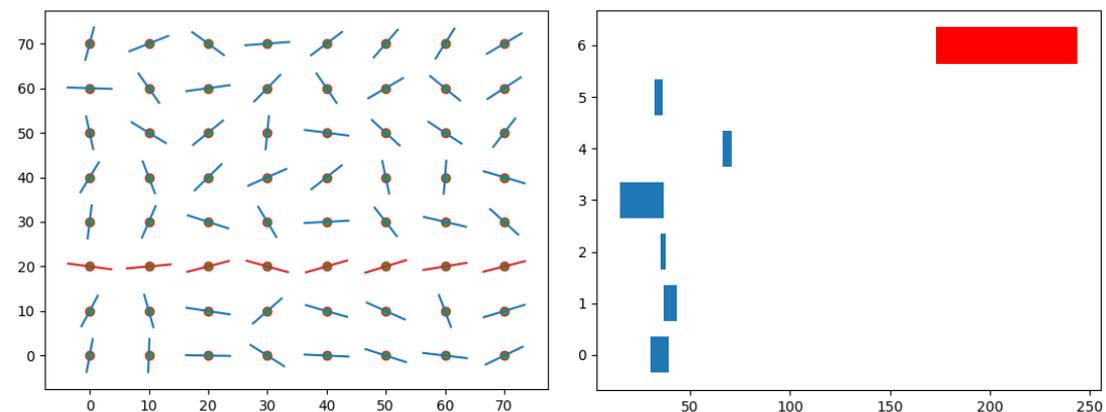


Рисунок 10. Эксперимент со средним уровнем шума.

Набор микросейсмических событий, не содержащий трещины, не образует значимых элементов первых гомологий (рис. 11).

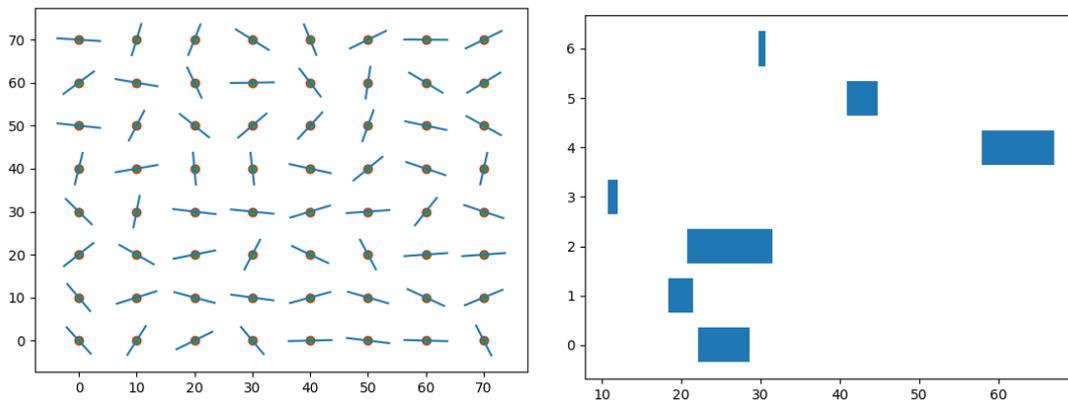


Рисунок 11. Эксперимент с высоким уровнем шума.

В результате подсчета длин продолжительности жизни элементов первых гомотопий мы можем наблюдать, что продолжительность жизни наиболее долгоживущего элемента из элементов первых гомотопий тем меньше, чем менее выражено наличие трещины в данных. Таким образом мы получили числовую характеристику, монотонно возрастающую с увеличением правдоподобности наличия трещины в данных.

ВЫВОДЫ

В работе дано сравнительно простое описание персистентных гомотопий – инструмента топологического анализа данных. Также представлено применение персистентных гомотопий, позволяющее в автоматическом режиме исследовать области пространства микросейсмических событий на наличие трещины, освобождая исследователя от необходимости глазами просматривать многомерные данные, допускающие большое количество интерпретаций, а также дающее возможность учитывать неточную и вероятностную природу детектирования микросейсмических событий.

Также представленный алгоритм предоставляет возможность естественного учета априорной информации и личных взглядов исследователя. Манипулируя метрикой, исследователь способен не только указать наиболее вероятное направление развития трещин, но и определить типы и параметры событий, которые он ожидает или не ожидает увидеть в результате.

Применение персистентных гомологий к задачам микросейсмического мониторинга раскрыто в статье не полностью, в дополнительных исследованиях нуждаются свойства метрик на пространствах микросейсмических событий и способы ввода априорной информации. Также представляет интерес интерпретация таких признаков баркодов одномерных гомологий, как время начала жизни наиболее долгоживущего элемента первых гомологий, которое может быть интерпретировано как наибольшее расстояния «разрыва» в правдоподобности узора микросейсмических событий, который интерпретируется как трещина или время окончания жизни наиболее долгоживущего элемента первых гомологий, которое, возможно, коррелирует с длиной трещины.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований и Кабинета министров Республики Татарстан в рамках проекта 18-47-160010.

СПИСОК ЛИТЕРАТУРЫ

1. *Herbert Edelsbrunner, John Harer. Persistent Homology — a Survey // Contemporary Mathematics. 2007. P. 103–120.*
2. *Herbert Edelsbrunner, Donald Kirkpatrick, Rainmund Siedel. On the shape of a set of points in the plane // IEEE Trans. Inform. Theory. 1983. P. 103–120.*
3. *Vanessa Robins. Toward computing homology from finite approximations // Topology Proceedings. 1999. V. 24, No. 1. P. 503–532.*
4. *Herbert Edelsbrunner, David Letscher, Afra Zomorodian. Topological Persistence and Simplification // Discrete Comput. Geom. 2002. V. 28. P. 511–533. DOI: 10.1007/s00454-002-2885-2.*
5. *David Cohen-Steiner, Herbert Edelsbrunner, John Harer. Stability of Persistence Diagrams // Discrete Comput. Geom. 2007. V. 37. P. 103–120. DOI: 10.1007/s00454-006-1276-5.*
6. *Xiaojin Zhu. Persistent Homology: An Introduction and a New Text Representation for Natural Language // IJCAI International Joint Conference on Artificial Intelligence. 2013. P. 1953–1959.*

7. *Gurjeet Singh, Facundo Memoli, Tigran Ishkhanov, Guillermo Sapiro, Gunnar Carlsson, Dario L. Ringach.* Topological analysis of population activity in visual cortex // *Journal of Vision.* 2008.V. 8, No. 11. DOI: 10.1167/8.8.11.
 8. *Vin de Silva, Robert Ghrist.* Homological sensor networks // *Notices of the American Mathematical Society.* 2007. V. 54, No. 1. P. 10–17.
 9. *Frederic Chazal, Steve Y. Oudot, Primoz Skraba, Leonidas J. Guibas.* Persistence-Based Clustering in Riemannian Manifolds // *Journal of the ACM.* 2013. V. 60, No. 6. Article No 14. DOI: 10.1145/1998196.1998212.
 10. *Sven Heydenreich, Benjamin Bruck, Joachim Harnois-Deraps.* Persistent homology in cosmic shear: constraining parameters with topological data analysis // *arXiv:2007.13724.*
 11. *Moon Duchin, Tom Needham, Thomas Weighill.* The (homological) persistence of gerrymandering // *arXiv:2007.02390.*
 12. *James Rutledge, Phillips W.S., House L.S., Zinno R.J.* Microseismic mapping of a Cotton Valley hydraulic fracture using Decimated downhole arrays // *SEG Technical Program Expanded Abstracts.* 1998. P. 338–341. DOI: 10.1190/1.1820422.
 13. *Serge A. Shapiro.* *Fluid-Induced Seismicity.* 2015. 289 p.
 14. *T.I. Urbancic, V. Shumila, J.T. Rutledge, R.J. Zinn.* Determining Hydraulic Fracture Behaviour Using Microseismicity // *Proceedings of 37th U.S. Rock Mech. Symp.* 1999. P. 991–996.
 15. *T.I. Urbancic, S.C. Maxwell.* Source Parameters of Hydraulic Fracture Induced Microseismicity // *SPE Annual Technical Conference and Exhibition, 2002.* DOI: 10.2118/77439-MS.
 16. *Tatjana Yanovskaya.* *Основы сейсмологии.* Изд-во С.-Пб. ун-та. 2008. 222 с.
 17. *Keiiti Aki, Paul G. Richards.* *Quantitative Seismology* // San Francisco, CA: Freeman, 1980. 49 p.
-

PERSISTENT HOMOLOGY: APPLICATION TO MONITORING HYDRAULIC FRACTURING

K. Y. Erofeev¹, M. T. Ziatdinov², E. V. Mokshin³

Kazan (Volga region) Federal University, Kazan

¹krllerof@gmail.com, ²gltronred@gmail.com, ³zhen-moks@yandex.ru

Abstract

Persistent homology is a topological data analysis tool which is reflecting changes in topological structure of data along its scale. Application of persistent homology to monitoring hydraulic fracturing which is allowing researchers to consider prior information in a natural way is given in the article.

Keywords: *topological data analysis, persistent homology, hydraulic fracturing, microseismic events.*

REFERENCES

1. *Herbert Edelsbrunner, John Harer.* Persistent Homology — a Survey // Contemporary Mathematics. 2007. P. 103–120.
2. *Herbert Edelsbrunner, Donald Kirkpatrick, Rainmund Siedel.* On the shape of a set of points in the plane // IEEE Trans. Inform. Theory. 1983. P. 103–120.
3. *Vanessa Robins.* Toward computing homology from finite approximations // Topology Proceedings. 1999. V. 24, No. 1. P. 503–532.
4. *Herbert Edelsbrunner, David Letscher, Afra Zomorodian.* Topological Persistence and Simplification // Discrete Comput. Geom. 2002. V. 28. P. 511–533. DOI: 10.1007/s00454-002-2885-2.
5. *David Cohen-Steiner, Herbert Edelsbrunner, John Harer.* Stability of Persistence Diagrams // Discrete Comput. Geom. 2007. V. 37. P. 103–120. DOI: 10.1007/s00454-006-1276-5.
6. *Xiaojin Zhu.* Persistent Homology: An Introduction and a New Text Representation for Natural Language // IJCAI International Joint Conference on Artificial Intelligence. 2013. P. 1953–1959.

7. *Gurjeet Singh, Facundo Memoli, Tigran Ishkhanov, Guillermo Sapiro, Gunnar Carlsson, Dario L. Ringach.* Topological analysis of population activity in visual cortex // *Journal of Vision.* 2008. V. 8, No. 11. DOI: 10.1167/8.8.11.
8. *Vin de Silva, Robert Ghrist.* Homological sensor networks // *Notices of the American Mathematical Society.* 2007. V. 54, No. 1. P. 10–17.
9. *Frederic Chazal, Steve Y. Oudot, Primoz Skraba, Leonidas J. Guibas.* Persistence-Based Clustering in Riemannian Manifolds // *Journal of the ACM.* 2013. V. 60, No. 6. Article No 14. DOI: 10.1145/1998196.1998212.
10. *Sven Heydenreich, Benjamin Bruck, Joachim Harnois-Deraps.* Persistent homology in cosmic shear: constraining parameters with topological data analysis // *arXiv:2007.13724.*
11. *Moon Duchin, Tom Needham, Thomas Weighill.* The (homological) persistence of gerrymandering // *arXiv:2007.02390.*
12. *James Rutledge, Phillips W.S., House L.S., Zinno R.J.* Microseismic mapping of a Cotton Valley hydraulic fracture using Decimated downhole arrays // *SEG Technical Program Expanded Abstracts.* 1998. P. 338–341. DOI: 10.1190/1.1820422.
13. *Serge A. Shapiro.* *Fluid-Induced Seismicity.* 2015. 289 p.
14. *T.I. Urbancic, V. Shumila, J.T. Rutledge, R.J. Zinn.* Determining Hydraulic Fracture Behaviour Using Microseismicity // *Proceedings of 37th U.S. Rock Mech. Symp.* 1999. P. 991–996.
15. *T.I. Urbancic, S.C. Maxwell.* Source Parameters of Hydraulic Fracture Induced Microseismicity // *SPE Annual Technical Conference and Exhibition, 2002.* DOI: 10.2118/77439-MS.
16. *Tatjana Yanovskaya.* *Основы сейсмологии.* Изд-во С.-Пб. ун-та. 2008. 222 с.
17. *Keiiti Aki, Paul G. Richards.* *Quantitative Seismology* // San Francisco, CA: Freeman, 1980. 49 p.

СВЕДЕНИЯ ОБ АВТОРАХ



ЕРОФЕЕВ Кирилл Юрьевич – аспирант, Казанский (Приволжский) федеральный университет, г. Казань.

Kirill Yurevich EROFEEV – postgraduate, Kazan (Volga region) Federal University, Kazan.

Email: krllerof@gmail.com



ЗИЯТДИНОВ Мансур Тагирович – научный сотрудник НИЛ «Квантовые методы обработки данных», Казанский (Приволжский) федеральный университет, г. Казань.

Mansur Tagirovich ZIATDINOV – researcher, "Quantum Information Science" Laboratory, Kazan (Volga region) Federal university, Kazan.

Email: gltronred@gmail.com



МОКШИН Евгений Владимирович – ведущий инженер исследователь, ООО Градиент-технологджи, г. Казань.

Evgenii Vladimirovich MOKSHIN – lead research engineer, Gradient Technology, Kazan.

Email: zhen-moks@yandex.ru

Материал поступил в редакцию 5 ноября 2020 года

УДК 004.514+004.928

ПРОСТРАНСТВЕННАЯ ОРИЕНТАЦИЯ ОБЪЕКТОВ НА ОСНОВЕ ОБРАБОТКИ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ ДЛЯ ГЕНЕРАЦИИ РАСКАДРОВОК

В. В. Кугуракова¹, Г. Ф. Сахибгареева², А. З. Нгуен³, А. М. Астафьев⁴

¹⁻⁴Казанский (Приволжский) федеральный университет, Высшая школа
информационных технологий и интеллектуальных систем

¹vlada.kugurakova@gmail.com, ²gulnara.sahibgareeva42@gmail.com,

³zung.nguyen.98@gmail.com, ⁴xiaomao472@gmail.com

Аннотация

Статья посвящена подходам в обработке текста на естественном языке для выяснения конкретного пространственного взаиморасположения объектов и трехмерной покадровой визуализации. Предложенные алгоритмы учитывают влияние явных ограничений извлеченных пространственных отношений и позволяют создавать вероятные трехмерные схемы расположения объектов на сцене. Такие пространственные схемы необходимы для уточнения воплощения замысла сценариста при проектировании видеоигр. Для создания удачных кадров использованы такие правила режиссуры, как учет плана, поворот камеры и другие композиционные нюансы.

Ключевые слова: преобразование текста в изображение, преобразование текста в сцену, преобразование текста в анимацию, понимание естественного языка, сценарий, видеоигра, сценарный прототип, положение камеры, режиссура.

ВВЕДЕНИЕ

Тема исследования связана с нашими попытками создать инструмент генерации раскадровок для сценарного прототипа видеоигры, которые являются черновым вариантом повествовательной составляющей, включая сценарий, внутриигровые тексты, инъекции предоставления возможностей для ведения диалога, повествовательные механики и другие инструменты передачи истории.

Архитектура разрабатываемого инструмента для генерации сценарного прототипа описана ранее в [1–3] и представлена на рис. 1.



Рисунок 1. Архитектура инструмента генерации сценарного прототипа.

Инструмент должен принимать на вход сценарий в виде текста на естественном языке. При этом мы считаем, что сценарий должен быть отформатирован по стандарту Голливуда [4] для лучшего распознавания сущностей.

Модуль редактирования должен давать возможность вносить изменения в сценарий, прототип и структуру сценария.

Модуль отображения должен реализовывать возможность просмотреть результат анализа документа: структура истории, сущности и статистика – количество персонажей, имена персонажей, реплики и т. д.

Генератор сценарного прототипа включает в себя три модуля, каждый из которых последовательно обрабатывает результат работы предыдущего:

1) модуль анализа текста должен распознавать необходимые сущности и связи между ними;

2) модуль анализа данных и построения структуры должен устанавливать целостную структуру сценария, связывать сущности между собой и строить структуру сценарного прототипа;

3) модуль визуализации должен подобрать необходимые трехмерные модели и распределить их по структуре сценарного прототипа; далее этот модуль должен отправить полученную визуализацию сценарного прототипа, а также визуализированную структуру сценария в Модуль воспроизведения.

Модуль воспроизведения должен позволять запускать получившийся сценарный прототип и диаграммы баланса (см. рис. 1).

В развитии идеи генерации сценарного прототипа [1] мы продвинулись на очередной этап – уточнение работы компонента визуализации. В данной работе мы представляем синтез прототипных решений для генерации трехмерной сцены для создания раскадровки. Для генерации используется текст, написанный на естественном языке.

Раскадровка – это промежуточный этап разработки генератора трехмерной анимации для интерактивного проекта. В данной работе рассматриваются такие нюансы генерации раскадровки, как постановка кадра и позиционирование предметов в пространстве.

Особое внимание посвящено пространственным предлогам, интерпретация которых помогает в генерации как общей трехмерной сцены, так и строения отдельного кадра.

Задачи пространственной ориентации объектов, создания трехмерной сцены и постановки кадра, решение которых мы предлагаем, включены в компонент визуализации.

СВЯЗАННЫЕ РАБОТЫ

Использование раскадровки в разработке игр сокращает время, деньги и другие ресурсы на создание компьютерной графики [2]. Более того, мы видим раскадровку как инструмент для тестирования повествовательной составляющей игры [1]. Компания Disney использует раскадровку как один из этапов разработки игр [5]. Функция генерации раскадровки по текстовому запросу есть в одной из версий Storyboarder, который мы использовали в качестве платформы для экспериментов [6].



(а) средний план;

(б) крупный план;

(в) очень крупный план

Рисунок 2. Примеры планов.

При генерации параметров для постановки кадра были выбраны следующие факты из практики кинорежиссуры ([7–11]): есть общий, средний и крупный планы, которые отличаются по количеству пространства, которое занимают в кадре ключевые объекты (см. рис. 2); есть 5 видов поворота камеры по оси X (с высоты птичьего полета, сверху вниз, на уровне глаз, снизу вверх, из-под ног) и 3 вида по оси Y (слева, спереди, справа), а также съемка со спины; расположение объектов регулируется правилами третей; камеры лучше двигать слева направо.

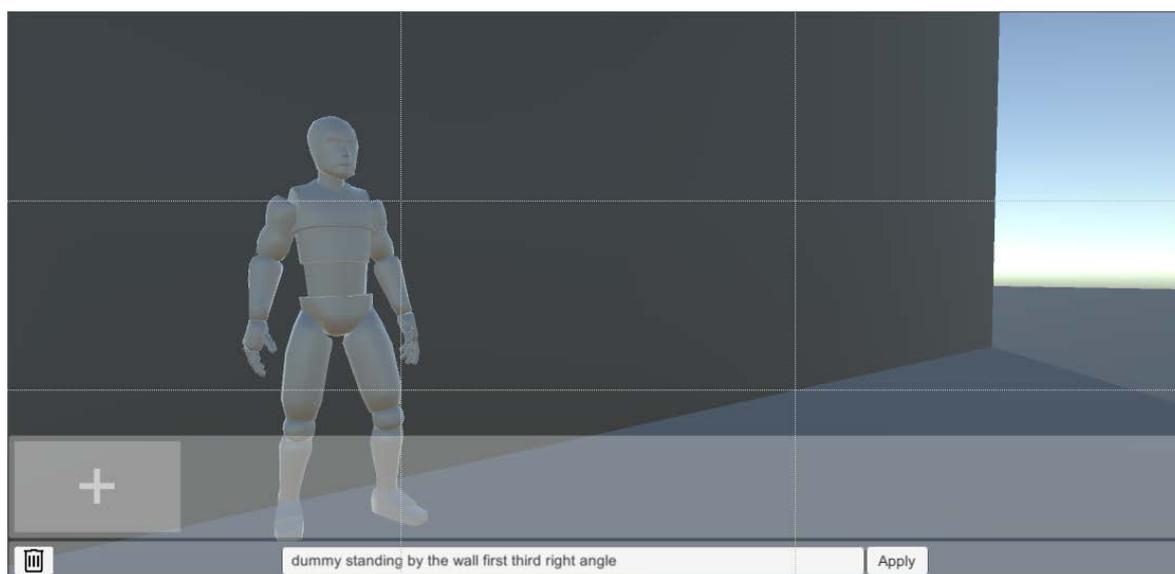


Рисунок 3. Изображение, разделенное по правилу третей.

Расположение объектов в соответствии с правилом третей – это принцип построения композиции, упрощающий правило золотого сечения. По этому принципу изображение следует рассматривать в виде девяти областей, которые разделены двумя линиями по вертикали и горизонтали, параллельными сторонам кадра, в определенных соотношениях. Важные части композиции обычно расположены на этих линиях или на их пересечениях, что акцентирует на них

внимание и преобладает над стандартным расположением объекта в центре кадра (см. рис. 2). Однако выбор тоже влияет на восприятие объектов в кадре. Направляющим фактором в данном случае является то, как люди воспринимают направление в кадре. Исходя из выводов исследования [12], считаем, что зрителям проще воспринимать движение слева направо, поэтому именно так и стоит располагать ключевые объекты. Однако бывают и менее тривиальные решения, где внимание зрителя сначала попадает в центр, затем – в правую, а затем – в левую части.

На разных этапах обработка текста происходит благодаря разным алгоритмам, это может быть система регулярных выражений, алгоритмы построения синтаксических и семантических сетей или обученные модели, но так или иначе важным является особенность цели интерпретации. В случае с генерацией раскадровки из постановки трехмерной сцены важными являются взаимное расположение объектов и камеры относительно друг друга, размеры локации и обстоятельства изменений в сцене.

В решении этой задачи помогает анализ пространственных предлогов, которые встречаются в английском тексте на естественном языке. Мы обратили внимание на следующие предлоги, которые помогают настроить сцену для статичных снимков: *by, in, to, with, nearby, between, front, behind, after, at, near, far, end, above, north, south, east, west, right, left, front*. Каждый предлог влияет на композицию объектов в кадре, а также на настройки камеры.

Естественный язык – простая и эффективная среда для описания визуальных идей и мысленных образов. Авторы [13], предвидя появление основанных на языке систем генерации трехмерных сцен, позволяющих обычным пользователям быстро создавать трехмерные сцены без необходимости изучать специальное программное обеспечение, еще в 2001 году представили собственный подход – *WordsEye* – систему для автоматического преобразования текста в репрезентативные 3D-сцены [14]. Интересно отметить и более поздние работы упомянутых авторов (см., например, [15; 16]), которые описывают детально процессы, необходимые для выяснения из текста относительного расположения объектов или уточнений месторасположения по описательным характеристикам.

Однако пространственное знание является важным аспектом мира и часто *не выражается явно* на естественном языке. WordsEye не решила проблему представления необъявленных фактов, которая является аналогом проблемы описания изображений (см., например, [17]). Эту проблему пытаются решать авторы [18] (кроме того, что извлекается представление пространственного знания из текста (слева, сверху и т. д.), изучается отображение между языком и геометрическими ограничениями, а также собирается статистика того, какие объекты встречаются в различных типах сцен, и рассматриваются их пространственные положения относительно друг друга.

НАША ИДЕЯ

Для получения удовлетворительного результата генерации мы разработали конвейер, который представлен в виде схемы (см. рис. 4).

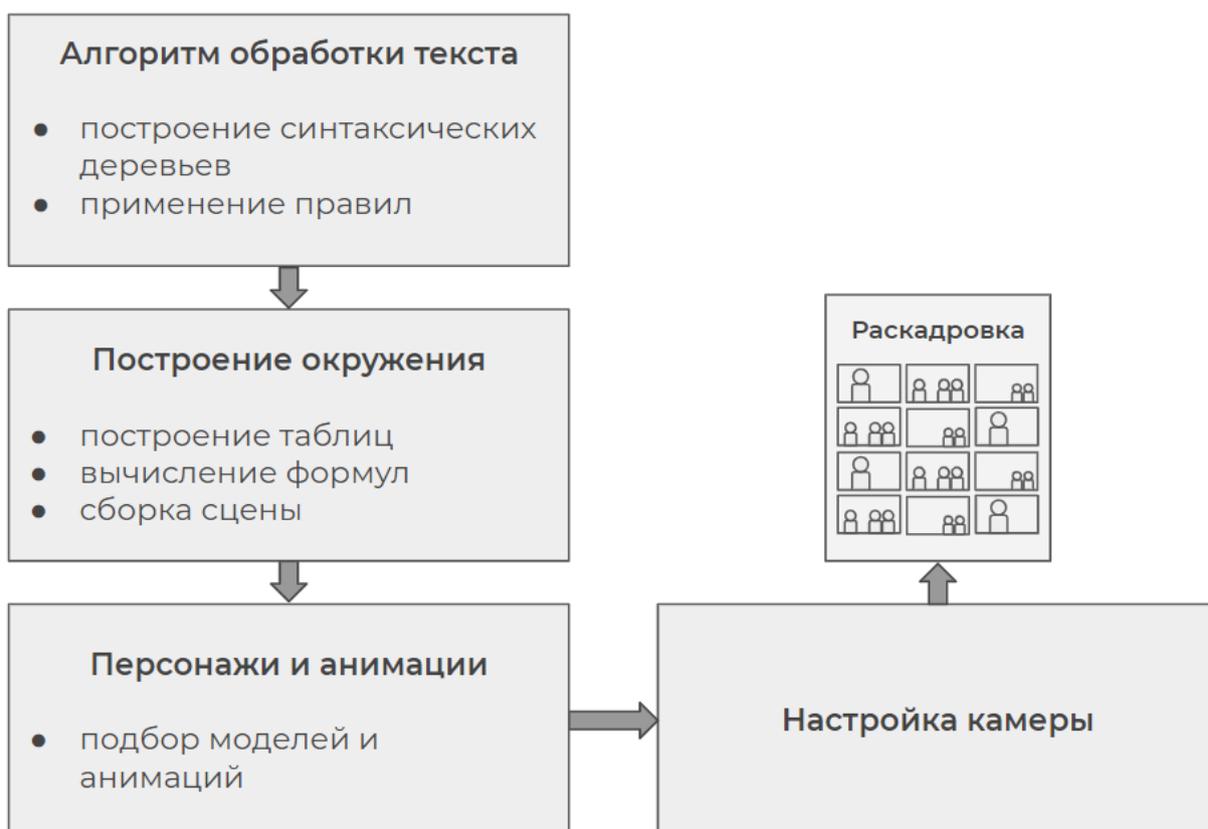


Рисунок 4. Схема конвейера.

Обработка текста. Большая часть кода разработана на python; кроме этого, используется C#.

Для извлечения информации из текста проведен анализ следующих биб-

лиотек: nltk, gensim, CoreNLP, OpenNLP, spaCy. В разработке использована библиотека spaCy [19], т. к. она производит вычисления быстрее других инструментов, кроме того, из её достоинств стоит упомянуть функционал для визуализации результата разбора фраз.

Сущностями, которые необходимо разметить в сцене, считаются имена существительные. Позы, в которые необходимо разместить персонажей, определяются по глаголам, относящимся к именам существительным. Относительное размещение сущностей определяется пространственными предлогами.

Алгоритм обработки для извлечения информации об окружении и предметах в сцене, настройках камеры выглядит так:

1. Маркировка тегов предметов на сцене;
2. Синтаксический разбор предложения;
3. Выделение сущностей, их состояний и (при наличии) их поз;
4. Определение пространственных предлогов;
5. Выделение информации о размещении сущности в кадре и на сцене, а также параметров настройки камеры;
6. Проверка наличия сочетаний слов, которые были установлены правилами, например, определение главного и второстепенного объектов;
7. Составление таблицы из извлеченных блоков текста.

Таблица 1. Взаимоотношения между объектами

word	distance	rotation_side	rotation_up	bool_is_interna	main_object
near	0	-1	-1	-1	-1
far	2	-1	-1	-1	-1
end	2	-1	-1	-1	-1
north	-1	90	-1	-1	-1
south	-1	180	-1	-1	-1
east	-1	0	-1	-1	-1
west	-1	270	-1	-1	-1
right	-1	0	-1	-1	-1
left	-1	180	-1	-1	-1
front	-1	90	-1	-1	-1
behind	-1	270	-1	-1	-1
above	-1	-1	180	-1	-1
under	-1	-1	0	-1	-1

Пространственные предлоги. Чтобы определить влияние пространствен-

ных предлогов на сущности, необходимо построить дерево взаимоотношений между объектами, что имеет смысл сделать при помощи выбранного синтаксического анализатора spaCy.

В конечном итоге должна быть составлена таблица (см. Табл. 1) следующего вида: предлог, зависимый или управляемый объект, угол поворота камеры, расстояние.

Машинное обучение. Метод, предложенный в [18], позволяет формировать выводы на основании статистических данных. Так можно получить значения конкретной дистанции между объектами в сцене.

Сборка сцены и режиссура. Далее автоматически построенная трехмерная сцена, учитывающая как полученные на основе анализа пространственных отношений в тексте (предлоги), так и необъявленные факты, в обычной речи следующие из здравого смысла, но которые необходимо сформировать заранее, так и на учете конкретных дистанций между объектами, превращаются в кадры, которые оформлены по правилам режиссерской и операторской работы (см. [7–11]).

ОПИСАНИЕ МЕТОДА

Как было сказано выше, для сценарного прототипа большое значение имеет функциональная визуализация, которую необходимо получать автоматически с возможностью редактирования. Генерация раскадровок – необходимый этап для перехода к анимации. К построению кадров раскадровки мы подходим через создание трехмерной сцены и последующий режиссерский выбор фокуса и положения камеры. Опишем поэтапно создание трехмерной сцены и создание раскадровки.

Создание трехмерной сцены происходит по следующему алгоритму:

1. На вход подается блок текста на естественном языке;
2. После его обработки формируются данные о сцене, описанной в тексте.

а. Для обработки текста была выбрана библиотека spaCy, которая привлекает к себе внимание своей скоростью работы и обладает всеми нужными инструментами и готовыми моделями для обработки естественного текста, в особенности английского: стоит отметить, что главным инструментом этой библиотеки

для фокуса нашей работы, а именно, определения относительного расположения объектов, стала возможность составить из текста дерево зависимостей. Приятным дополнением `displaCy` является возможность рассмотреть, как был разобран текст, и увидеть взаимосвязь между словами в предложении (см. рис. 5). Поскольку синтаксическое дерево предоставляет еще и POS-тэги для слов, а также именованные связи, то основная идея заключается в использовании этих связей для составления правил, по которым будет выделяться нужная информация.

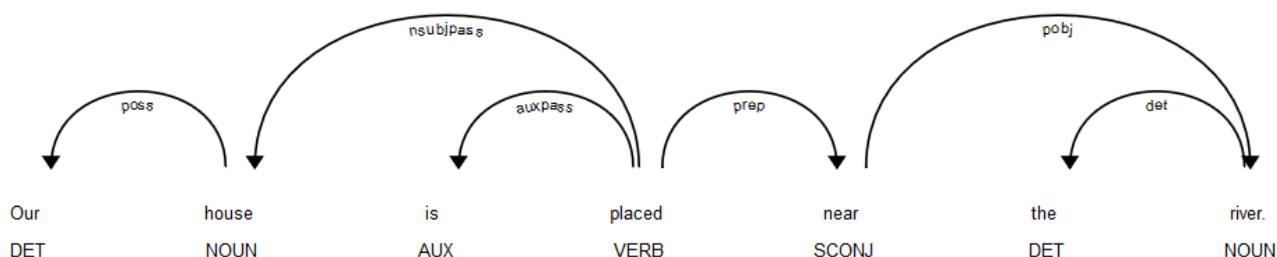


Рисунок 5. Утилита `displaCy` из библиотеки `sraCy` позволяет визуализировать взаимосвязи между словами в предложении.

b. В английском языке для обозначения позиции объекта относительно другого объекта используются пространственные предлоги. Рассмотрим пример, представленный рис. 6. Здесь родителем интересующего нас предлога «near» является глагол «to place», а существительные «house» и «river» являются потомками глагола и предлога соответственно. И в данном случае главным существительным будет «house», а второстепенным – «river». Таким образом можно создать правило, которое будет извлекать блоки текста, если в предложении появятся 2 существительных, глагол и предлог, связанные между собой определенным способом.

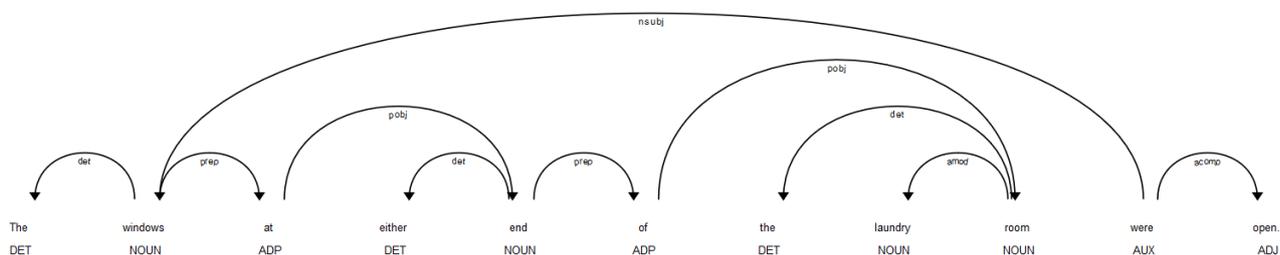


Рисунок 6. Пример взаимосвязей в предложении.

c. Следует отметить, что пространственные определения, такие как «near the river», можно ставить не только после глагола, но и после существительного. В

данном случае на самом деле просто опускается глагол «to be» для более красивой речи.

d. Если рассмотреть антоним предлогу «near», то можно заметить, что «far» всегда используется в связке с «from» (у «near» часто опускается «to»), образуя тем самым еще одно правило.

e. Иногда указанием на относительное месторасположение объектов служат не пространственный предлог, а существительное. Тогда эти два существительных связаны друг с другом сочетанием «at end of», состоящим из двух предлогов и существительного. Здесь существительное «end» играет ключевую роль, обозначая, что окно находится в конце комнаты, а два предлога не несут никакой смысловой нагрузки и стоят лишь в соответствии с правилами английского языка.

f. Немаловажно в данных правилах определить, какое существительное из двух будет главным существительным, а какое второстепенным. Если, например, в предложении «стол рядом со стулом» поменять местами существительные, то сути это не изменяет, однако если поменять в «стол справа от стула», то уже возникнет ошибка. Все изменяется в зависимости от того, относительно которого объекта происходит сравнение. В данном случае и в русском, и в английском языках главным объектом всегда будет существительное, находящееся в уточняющем дополнении.

g. Поскольку не все предлоги являются пространственными и не все существительные передают информацию о взаимном месторасположении объектов, то было принято решение использовать словарь. Благодаря ему, обрабатывая текст, инструмент будет исключать те блоки предложения, в которых не встречается нужные слова. Также, поскольку в последствии потребуются расставить объекты на сцене, за каждым словом были закреплены некоторые числовые значения.

3. После создания дерева зависимостей библиотекой spaCy выделяются все существительные, которые есть в предложениях. Чтобы сопоставлять их с «объектами мира», каждый из них лемматизируется (то есть приводится в словарную форму). После этого проверяется наличие сочетаний слов, которые были установлены правилами, созданными ранее. Из извлеченных блоков текста со-

ставляется таблица с полями:

- a. Главный объект;
- b. Второстепенный объект;
- c. Взаимное месторасположение объектов;
- d. Числовые значения – удаленность одного объекта от другого и относительный угол.

4. Затем конкретизируются параметры о месторасположении объектов относительно друг друга.

a. Параметр Rotation содержит информацию о позиции объекта относительно другого объекта (справа, слева, сверху, снизу). В качестве численного параметра используются градусы (см. рис. 7). При размещении относительно не главного, а второстепенного объекта, значение пересчитывается ($Rotation - 180$).

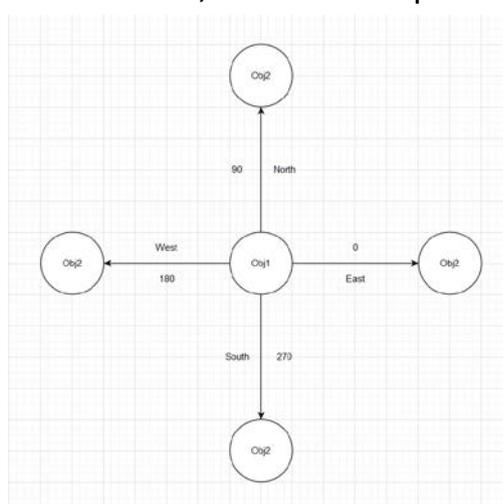


Рисунок 7. Градусы для обозначения позиции объекта относительно другого объекта.

5. Параметр Position (см. рис. 8) содержит информацию о расстоянии между объектами (от 0 до 2). Это обусловлено тем, что такие слова, как «близко» и «далеко», являются довольно расплывчатым определением расстояния между объектами. Человек воспринимает информацию, опираясь на свой накопленный опыт, поэтому для него «корабль около порта» воспринимается буквально, что корабль находится непосредственно у порта, а «человек у горы» – что человек находится либо у подножия горы, либо в двух трех километрах от нее.

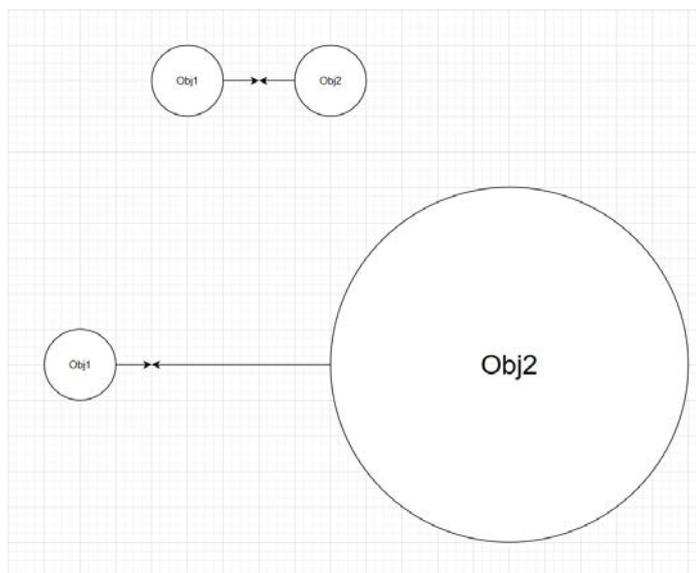


Рисунок 8. Расстояние между объектами задается радиусом.

6. На основании сформированных параметров создается карта местности с расположением объектов (см. [20]).

7. В работе [18] решают проблему уточнения числовых значений.

При генерации трехмерной сцены на основе текста на естественном языке важно не просто построить дерево зависимостей извлеченных сущностей, но и достроить описание сцены, чтобы на выходе получить адекватный результат. Когда на вход поступает запрос вроде «В комнате стол с тортом», на выходе должен получиться торт, который стоит на столе, а не под ним на полу. Такое уточнение дает модель, которую разработали в 2014 году [18], и она показала хорошие результаты (рис. 9).

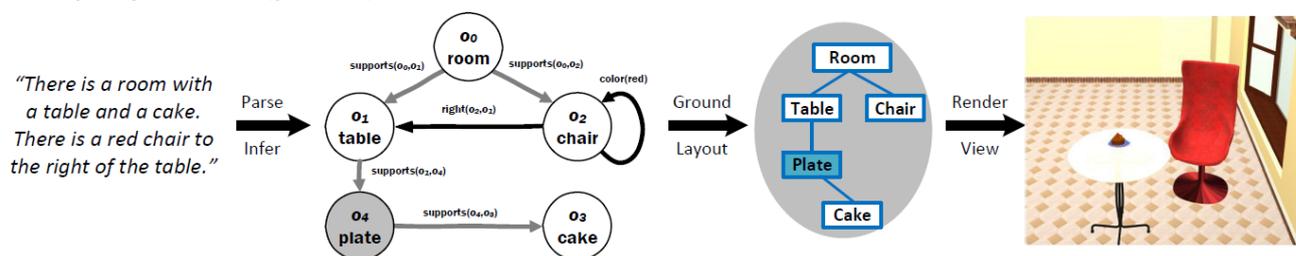


Рисунок 9. Конвейер генерации трехмерной сцены (по [18]).

Для обработки текстов в инструменте используется Stanford CoreNLP. В результате анализа текста происходит построение геометрического графа, в вершинах которого располагаются именованные трехмерных объектов, а ребрами являются пространственные отношения между ними. Пространственные отно-

шения определяются шаблонами. Чтобы получить их, в системе используются условная вероятность и правило Байеса. Полученные правила вносят изменения в граф, построенный ранее.

Кроме этого, модель активно обучается в процессе работы пользователя в редакторе полученной трехмерной сцены. Однако в [21] утверждают, что WordsEye превосходит разработку [18] с точки зрения понимания языка и объема репозитория трехмерных моделей.

Так как необходимо извлекать из текста неявную информацию, то безусловно нужно включить один из упомянутых в [18] подходов или иметь возможность выбора наилучшего решения из множества – как минимум, для уточнения вертикального позиционирования объектов, обрабатывая положения, формируемые соответствующими предлогами (ср. русские предлоги *с, над, под, на, при, у, около, за, перед, из-за, из-под* и т. п.).

Создание раскадровок

8. Во время разбора текста для получения общей трехмерной сцены с детальной проработкой расстановки объектов на ней параллельно происходит выделение тегов, как описывающих сцену, так и описывающих влияние на параметры постановки камеры. Важнейшей задачей являются обеспечение условий для постановки различных вариантов композиций и изменение состояний персонажей и объектов, присутствующих на собранной сцене.

Разработка возможностей, связанных с построением сцен и раскадровки для динамического действия, не является приоритетной, несмотря на некоторые сходства ее производства, значительно уступающих существующим различиям. Взаимодействие анимированных персонажей друг с другом относится к этим различиям, так как конкретная анимация персонажей за редким исключением задает композицию кадра. Однако в случаях, когда важно именно само взаимодействие персонажей, его проработка выходит за рамки задач раскадровки.

9. После синтаксического разбора предложений полученное формализованное описание включает конкретные предметы сцены, их состояния, пространственные отношения, а также информацию, необходимую для размещения объекта в кадре и на сцене.

10. По найденным тегам вызываются соответствующие методы, кото-

рые ищут группу объектов, относительно которой размещается субъект пространственного предлога. В разных предлогах разные приоритеты объекта и субъекта в кадре, взаимное расположение объектов. Помимо координат объектов, подлежащих изменениям на сцене, изменяется также их значение поворота.

11. Далее относительно готовой трехмерной сцены специальный компонент изменяет положение камеры в ней, исходя из заданных параметров. Параметры кадра соответствуют выведенным ранее типам съемки и различным углам поворота камеры. Их генерация происходит посредством компонента, который обрабатывает полученный результат разбора предложения, учитывая количество важных частей речи, конкретное влияние некоторых из них на тип кадра, количество основ предложения. Однако полученные таким образом параметры съемки уступают по приоритету заданным вручную, которые также собираются с помощью этого компонента из текстового описания.

12. Собранные таким образом параметры влияют на вычисление стартовой позиции камеры. Постановка камеры начинается со стартовой позиции, которая вычисляется, исходя из положения ключевых объектов таким образом, что все они находятся в поле зрения камеры. Далее это положение изменяется в зависимости от параметров. Исходя из полученных параметров, разными путями возможно дальнейшее редактирование положения и угла наклона камеры.



Рисунок 10. Интерфейс инструмента, выбран кадр для редактирования.

Когда постановка сцены и камеры завершена, становится доступна запись всех параметров данного кадра. Как только кадр устраивает пользователя, он

может приступить к созданию следующего, который будет начат с тех же объектов и типов съемки, что и предыдущий. Последовательность полученных кадров доступна для редактирования инструментом по нажатию на кадр. Также изменения возможны и путем прямого введения тегов для описания сцены (см. рис. 10).

Последовательно обрабатывая текстовое описание сцен, пользователь получает в конце работы последовательность раскадровки, представленную в виде графических растровых изображений, построенных на основе коллекции трехмерных моделей, с которыми ассоциированы извлекаемые теги-наименования объектов или персонажей (см. [22]).

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Рассмотрим пример использования разработанного инструмента и сравнение результатов работы с использованными в фильме материалами. Для этого используем вариант словесного описания, происходящего на изображениях (см. рис. 11) на основе материалов киносериала «Мост» [23].



Рисунок 11. Сценарная раскадровка и итоговые кадры.

Первый в последовательности кадр удалось получить из описания “<character> is standing firmly by the cliff first third”. Таким образом, данные об

угле поворота камеры и типе съемки были получены, исходя из наличия в тексте наречий или дополнений, а также глагола.

Второй кадр был получен из “<character1> and <character2> are talking with <character3> by the cliff right angle”. По аналогии с первым, часть информации композиции кадра была получена из описания, но также была изменена с помощью добавления в текст конкретных тегирующих фраз, влияющих на композицию.



Рисунок 12. Полученная последовательность кадров.

Третий кадр был получен из описания: “<character1> and <character2> are staring at field backshot last third right angle”. Несмотря на значительные рукотворные изменения композиции, результат оказался лишь частично схож с оригиналом из-за различий в рельефе сцены. Таким образом, с применением описания и добавления некоторых тегов типа съемки получается последовательность кадров, в большей степени соответствующая рассмотренному примеру (рис. 12).

Также с помощью инструмента удалось воспроизвести стандартную диалоговую последовательность (см. рис. 13) – когда камера переключается с лица на лицо во время диалога.



Рисунок 13. Диалоговая последовательность.

Исходя из приведенных выше примеров, можно сделать вывод о том, что выбранные подходы сборки сцены и конкретная реализация позволяют получить ценный результат при работе с прототипированием сценария в виде раскадровки. Наиболее полезен этот инструмент будет в работе именно со статичными компоновками.

ДАЛЬНЕЙШЕЕ РАЗВИТИЕ

Описываемый функционал инструмента генерации сценарного прототипа ещё имеет массу чёрных пятен, ещё только начаты исследования следующих функций:

- Создание реалистичного окружения – для этого необходимы распознавание элементов ландшафта (реки, озёра, дороги) и их размещение относительно точки зрения повествующего.
- Визуализация сюжетных линий и реализация вариативности сюжетов.
- Автоматическое проектирование сценария с разбивкой на отдельные сцены и кадры.
- Автоматическая примитивизация и автоматическое тегирование публичных трехмерных моделей для увеличения архива моделей, используемых при генерации раскадровок. Сложность такого подхода связана также с вынесением данных о модели из её бинарного представления, а информация о подобном декодировании не всегда доступна.
- Автоматический риггинг⁸ модели персонажа для получения состояния по ключевым фразам, упоминающим некоторые характеристики: телосложение, пол, действие и т. п.

⁸ Риггинг – расстановка вершин на меш (каркас) трехмерной модели для последующего создания анимации.

- Использование тегов для разметки эмоционального различия анимаций персонажей.
- Генерация светового решения сцены соответственно эмоциональному посылу действия и композиции кадра.

Конечно, важна и собственно реализация многокомпонентной архитектуры, которая подразумевает заменяемость функционала отдельных элементов инструмента генерации сценарного прототипа.

ЗАКЛЮЧЕНИЕ

Разработанный подход позволяет строить быстрые (относительно стандартного процесса создания раскадровок) визуализации на основе генерации трехмерной сцены, для чего были созданы правила, которые позволяют извлекать полезную информацию из текста и формировать итоговый кадр по тегам, описывающим положение камеры. Тестовая генерация раскадровки оказалась быстрее ручной сборки в 6 раз.

Таким образом, представлена основная часть инструмента для генерации сценарных и игровых прототипов [24] – как для компьютерных игр, так и для serious games или образовательных тренажеров (см., например, [25]).

СПИСОК ЛИТЕРАТУРЫ

1. Сахибгареева Г.Ф., Бедрин О.А., Кугуракова В.В. Разработка компонента генерации визуализации сценарного прототипа видеоигр // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции. М.: ИПМ им. М.В. Келдыша, 2020. С. 581–603.

2. Сахибгареева Г.Ф. Разработка решения для визуализации сценарного прототипа инструментами генерации раскадровок // Выпускная квалификационная работа магистратуры Высшей школы ИТИС. Казанский федеральный университет, Казань, 2020. 75 с.

URL: https://kpfu.ru/student_diplom/10.160.178.20_X87SBS2R8OQUD822Q1Y9MOGMYD_6QTSVCTR0JTXAC5Z54A_W6A_Sahibgareeva.pdf.

3. Сахибгареева Г.Ф. Разработка инструмента для создания игрового сценарного прототипа: выпускная квалификационная работа бакалавриата Высшей школы ИТИС. Казанский федеральный университет, Казань, 2018. 57 с.
URL: https://kpfu.ru/student_diplom/10.160.178.20_9311629_F_Sahibgareeva.pdf.
4. Хейг М. Голливудский стандарт: как написать сценарий для кино и ТВ, который купят / Майкл Хейг; Пер. с англ. М.: Альпина нон-фикшн, 2017. 388 с.
5. Lee N., Madej K. Disney Stories: Getting to Digital. Springer, 2012. 114 с.
6. Storyboarder: Wonder Unit. Режим доступа:
URL: <https://wonderunit.com/storyboarder/> (Дата обращения: 12.05.2020).
7. Thompson R. Grammar of the Shot. Focal Press, 1998. 219 p.
8. Hockrow R. Out of Order: Storytelling Techniques for Video and Cinema Editors. Peachpit Press, 2014. 240 p.
9. Dise J. Filmmaking 101: Camera Shot // Explora, 2017. Режим доступа:
URL: <https://www.bhphotovideo.com/explora/video/tips-and-solutions/filmmaking-101-camera-shot-types> (Дата обращения: 6.04.2020).
10. Gallea R., Ardizzone E., Pirrone R. Automatic aesthetic photo composition // International Conference on Image Analysis and Processing. Springer, Berlin, Heidelberg, 2013. P. 21–30.
11. Olivier P., Tsujii J. A computational view of the cognitive semantics of spatial prepositions // Proceedings of the 32nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, 1994. P. 303–309.
12. Egizii M.L. et al. Which way did he go? Directionality of film character and camera movement and subsequent spectator interpretation // International Communication Association conference. Phoenix, AZ, 2012. Режим доступа:
URL: https://www.researchgate.net/publication/228448619_Which_Way_Did_He_Go_Directionality_of_Film_Character_and_Camera_Movement_and_Subsequent_Spectator_Interpretation (Дата обращения: 28.04.2020).
13. Coyne B., Sproat R. WordsEye: an automatic text-to-scene conversion system // Proceedings of the 28th annual Conference on Computer Graphics and Interactive Techniques, 2001. P. 487–496.

14. *Ulinski M., Coyne B., Hirschberg J.* Evaluating the WordsEye Text-to-Scene System: Imaginative and Realistic Sentences // Proceedings of the Eleventh International Conference on Language Resources and Evaluation, 2018. P. 1493–1499.
 15. *Coyne B., Sproat R., Hirschberg J.* Spatial relations in text-to-scene conversion // Computational Models of Spatial Language Interpretation, Workshop at Spatial Cognition, 2010. 8 p.
 16. *Coyne B., Klapheke A., Rouhizadeh M., Sproat R., Bauer D.* Annotation tools and knowledge representation for a text-to-scene system // Proceedings of COLING 2012: Technical Papers. 2012. P. 679–694.
 17. *Elliott D., Keller F.* Image description using visual dependency representations // Proceedings of Empirical Methods in Natural Language Processing (EMNLP). 2013. P. 1292–1302.
 18. *Chang A.X., Savva M., Manning Ch.D.* Learning Spatial Knowledge for Text to 3D Scene Generation // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014. P. 2028–2038.
 19. spaCy: библиотека для NLP // Официальный сайт spaCy с документацией. Режим доступа: URL: <https://spacy.io> (Дата обращения: 10.06.2020).
 20. *Нгуен А.З.* Генерация окружения на основе текстового описания: выпускная квалификационная работа бакалавриата Высшей школы ИТИС. Казанский федеральный университет, Казань, 2020. 45 с.
URL: https://kpfu.ru/student_diplom/10.160.178.20_OW9ARNOUY_98UU56A_D0JDQMEBFOF__66GZP8OQRFCLC4RY26N_F_Nguyen.pdf.
 21. *Hassani K., Lee W.-S.* Visualizing Natural Language Descriptions // ACM Computing Surveys. 2016. V. 49(1). P. 1–34.
 22. *Астафьев А.М.* Разработка инструмента для сборки сцен по тегам: выпускная квалификационная работа бакалавриата Высшей школы ИТИС. Казанский федеральный университет, Казань, 2020. 75 с.
URL: https://kpfu.ru/student_diplom/10.160.178.20_FPEBER9KDIZQVYJAE3VRTIFYWZV_CDDM972OPP2I28S0EEFABT_Astafev.pdf.
 23. Episodes of “The Bridge” / Киносериял // BBC. 2020. Режим доступа: URL: <https://www.bbc.co.uk/programmes/b01gxlxj/episodes/guide> (Дата обращения: 24.04.2020).
-

24. Сахибгареева Г.Ф., Кугуракова В.В. Концепт инструмента автоматического создания сценарного прототипа компьютерной игры // *Электронные библиотеки*. 2018. Т. 21, №3-4. С. 235–249.

25. Антонов И.О., Зезегова К.В., Кугуракова В.В., Лазарев Е.Н., Хафизов М.Р. Программирование запахов для виртуального осмотра места происшествия // *Электронные библиотеки*. 2018. Т. 21, № 3-4. С. 301–313.

SPATIAL ORIENTATION OF OBJECTS BASED ON PROCESSING OF A NATURAL LANGUAGE TEXT FOR STORYBOARD GENERATION

V. V. Kugurakova¹, G. F. Sahibgareeva², A. Z. Nguyen³, A. M. Astafiev⁴

¹⁻³ *Kazan Federal University (KFU), Institute of Information Technology and Intelligent Systems (ITIS)*

¹vlada.kugurakova@gmail.com, ²gulnara.sahibgareeva42@gmail.com,

³zung.nguyen.98@gmail.com, ⁴xiaomao472@gmail.com

Abstract

The article is devoted to our approaches to processing text in natural language to clarify the specific spatial relationship of objects and three-dimensional frame-by-frame visualization. The proposed approach allows us to show how the explicit constraints of the extracted spatial relationships affect and makes it possible to create possible layouts of objects on the scene. Natural language interpretations for spatial knowledge can generate three-dimensional scenes, which in turn are necessary to translate the scriptwriter's intent into the design of video games. The work also takes into account the rules of directing to create successful shots. Among them, accounting for the plan, camera rotation, as well as compositional nuances.

Keywords: text-to-picture conversion, text-to-scene conversion, text-to-animation conversion, natural language understanding, video game, scenario, scenario prototype.

REFERENCES

1. *Sahibgareeva G.F., Kugurakova V.V., Bedrin O.A.* Razrabotka resheniya dlya vizualizacii scenarnogo prototipa instrumentami generacii raskadrovok // Vserossijskaya nauchnaya konferenciya "Nauchnyj servis v seti Internet". 2020. S. 581–603.
2. *Sahibgareeva G.F.* Razrabotka resheniya dlya vizualizacii scenarnogo prototipa instrumentami generacii raskadrovok // Vypusknaya kvalifikacionnaya rabota magistratury Vysšej shkoly ITIS. Kazanskij federal'nyj universitet, Kazan', 2020. 75 s.
URL: https://kpfu.ru/student_diplom/10.160.178.20_X87SBS2R8OQUD822Q1Y9MOGMYD_6QTSVCTROJTXAC5Z54A_W6A_Sahibgareeva.pdf.
3. *Sahibgareeva G.F.* Razrabotka instrumenta dlya sozdaniya igrovogo scenarnogo prototipa: vypusknaya kvalifikacionnaya rabota bakalavriata Vysšej shkoly ITIS. Kazanskij federal'nyj universitet, Kazan', 2018. 57 s. URL: https://kpfu.ru/student_diplom/10.160.178.20_9311629_F_Sahibgareeva.pdf.
4. *Heĭg M.* Gollivudskiĭ standart: Kak napisat' scenariĭ dlya kino i TV, kotoryĭ kupyat / Maĭkl Heĭg; Per. s angl. M.: Al'pina non-fikshn, 2017. 388 s.
5. *Lee N., Madej K.* Disney Stories: Getting to Digital. Springer. 2012. 114 p.
6. Storyboarder: Wonder Unit. URL: <https://wonderunit.com/storyboarder/> (Date of access: 12.05.2020).
7. *Thompson R.* Grammar of the Shot. Focal Press. 1998. 219 p.
8. *Hockrow R.* Out of Order: Storytelling Techniques for Video and Cinema Editors. Peachpit Press. 2014. 240 p.
9. *Dise J.* Filmmaking 101: Camera Shot // Explora. 2017. URL: <https://www.bhphotovideo.com/explora/video/tips-and-solutions/filmmaking-101-camera-shot-types> (Date of access: 6.04.2020).
10. *Gallea R., Ardizzone E., Pirrone R.* Automatic aesthetic photo composition // International Conference on Image Analysis and Processing. Springer, Berlin, Heidelberg. 2013. P. 21–30.
11. *Olivier P., Tsujii J.* A computational view of the cognitive semantics of spatial prepositions // Proceedings of the 32nd annual meeting on Association for

Computational Linguistics. Association for Computational Linguistics. 1994. P. 303–309.

12. *Egizii M.L. et al.* Which way did he go? Directionality of film character and camera movement and subsequent spectator interpretation // International Communication Association Conference. Phoenix, AZ, 2012.

URL: https://www.researchgate.net/publication/228448619_Which_Way_Did_He_Go_Directionality_of_Film_Character_and_Camera_Movement_and_Subsequent_Spectator_Interpretation (Date of access: 28.04.2020).

13. *Coyne B., Sproat R.* WordsEye: an automatic text-to-scene conversion system // Proceedings of the 28th annual conference on Computer graphics and interactive techniques. 2001. P. 487–496.

14. *Ulinski M., Coyne B., Hirschberg J.* Evaluating the WordsEye Text-to-Scene System: Imaginative and Realistic Sentences // Proceedings of the Eleventh International Conference on Language Resources and Evaluation. 2018. P. 1493–1499.

15. *Coyne B., Sproat R., Hirschberg J.* Spatial relations in text-to-scene conversion // Computational Models of Spatial Language Interpretation, Workshop at Spatial Cognition. 2010. 8 p.

16. *Coyne B., Klapheke A., Rouhizadeh M., Sproat R., Bauer D.* Annotation tools and knowledge representation for a text-to-scene system // Proceedings of COLING 2012: Technical Papers. 2012. P. 679–694.

17. *Elliott D., Keller F.* Image description using visual dependency representations // Proceedings of Empirical Methods in Natural Language Processing (EMNLP). 2013. P. 1292–1302.

18. *Chang A. X., Savva M., Manning Ch.D.* Learning Spatial Knowledge for Text to 3D Scene Generation // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014. P. 2028–2038.

19. spaCy: библиотека для NLP // Официальный сайт spaCy с документацией. URL: <https://spacy.io> (Date of access: 10.06.2020).

20. *Nguen A.Z.* Generaciya okruzeniya na osnove tekstovogo opisaniya: vypusknaya kvalifikacionnaya rabota bakalavriata Vyshej shkoly ITIS. Kazanskij federal'nyj universitet, Kazan', 2020. 45 s.

URL: https://kpfu.ru/student_diplom/10.160.178.20_OW9ARNOUY_98UU56A_DOJDQMEBFOF__66GZP8OQRFCLC4RY26N_F_Nguyen.pdf.

21. *Hassani K., Lee W.-S.* Visualizing Natural Language Descriptions // ACM Computing Surveys. 2016. V. 49(1). P. 1–34.

22. *Astaf'ev A.M.* Razrabotka instrumenta dlya sborki scen po tegam: vypusknaya kvalifikacionnaya rabota bakalavriata Vysšej shkoly ITIS. Kazanskij federal'nyj universitet, Kazan', 2020. 75 s.

URL: https://kpfu.ru/student_diplom/10.160.178.20_FPEBER9KDIZQVYJAE3VRTIFYWZB_CDDM972OPP2I28S0EEFABT_Astafev.pdf.

23. Episodes of “The Bridge” / movie series // BBC. 2020.

URL: <https://www.bbc.co.uk/programmes/b01gxlxj/episodes/guide> (Date of access: 24.04.2020).

24. *Sahibgareeva G.F., Kugurakova V.V.* Koncept instrumenta avtomaticheskogo sozdaniya scenarnogo prototipa komp'yuternoj igry // Elektronnye biblioteki. 2018. T. 21. No. 3-4. S. 235–249.

25. *Antonov I.O., Zezegova K.V., Kugurakova V.V., Lazarev E.N., Hafizov M.R.* Programmirovaniye zapahov dlya virtual'nogo osmotra mesta proisshestiya // Elektronnye biblioteki. 2018. T. 21, No. 3-4. S. 301–313.

СВЕДЕНИЯ ОБ АВТОРАХ



КУГУРАКОВА Влада Владимировна – к. т. н., доцент кафедры программной инженерии ИТИС КФУ, руководитель НИЛ разработки AR/VR приложений и компьютерных игр. Сфера научных интересов – генерация реалистичной визуализации трехмерных сцен и различные аспекты проектирования игр.

Vlada Vladimirovna KUGURAKOVA, PhD., Docent of ITIS KFU, Head of Laboratory «AR/VR applications and Gamedev». Research interests include different approaches in game design.

email: vlada.kugurakova@gmail.com.



САХИБГАРЕЕВА Гульнара Фаритовна – ассистент кафедры программной инженерии ИТИС КФУ. Сфера научных интересов – игровая сценаристика, нарративный дизайн, изучение вопроса эффективности создания сценарного прототипа и возможности автоматизировать данный процесс.

Gulnara Faritovna SAHIBGAREEVA – assistant of ITIS KFU. Sphere of research: game scripting, narrative design, studying the issue of the effectiveness of creating a scenario prototype and the ability to automate this process.

email: gulnara.sahibgareeva42@gmail.com



НГУЕН Ань Зунг – выпускник бакалавриата ИТИС КФУ. Сфера научных интересов – использование нейронных сетей в разработке игр.

An' Zung NGUYEN – Bachelor of the ITIS KFU. Research interests include neural networks for game development.

email: zung.nguyen.98@gmail.com



АСТАФЬЕВ Андрей Максимович – студент магистратуры ИТИС КФУ. Сфера интересов: разработка игр и AR/VR приложений на Unity.

Andrey Maksimovich ASTAFIEV – Masters student of ITIS KFU. Research interests include development in Unity.

Email: xiaomao472@gmail.com

Материал поступил в редакцию 8 ноября 2020 года

УДК 004.91

ПРОГРАММА «ИСТОРИЯ ГЕНИАЛЬНОГО ОТКРЫТИЯ»

Р. В. Мосолов

*Казанский (Приволжский) федеральный университет, Казань;
Schematus (USA), Россия, 420015, г. Казань, ул. Пушкина, д. 54/1 (коворкинг
«Телеграф Space»)*

R.V.Mosolov@ya.ru

Аннотация

Настоящая статья описывает концепцию программного обеспечения (ПО) «История гениального открытия», имеющего ряд сходств с программой GitHub, получившей широкую известность в профессиональном сообществе программистов. Программа призвана решать две основные научные проблемы: сохранять научно-культурное наследие российских учёных и аккумулировать первичные данные, позволяющие количественно измерить тенденции становления научных теорий, тем самым дополнив концепцию «научных революций» Т. Куна. Программа позволит сохранять исторически значимые научные достижения, минимизируя вероятность их бесследных потерь вследствие преждевременно ухода учёных из жизни. Идея разработки программы базируется на пересечении пяти научных направлений – программной инженерии, социологии, философии, права и истории – и появилась в стенах Казанского (Приволжского) федерального университета при изучении Big Data Science.

Ключевые слова: *История гениального открытия, научное наследие, культурное наследие, закономерности гениальности, ПО для учёных, программа для учёных, GitHub для учёных.*

ВВЕДЕНИЕ

Середина XX в. в США и конец XX в. в России ознаменовали собой периоды расцвета компьютерных технологий. Сегодня глобальную коммуникацию между людьми сложно представить осуществимой без наличия соответствующих гаджетов (смартфонов, компьютеров, планшетов и др.). Однако польза от инфор-

мационных технологий может быть получена не только широкими социальными слоями, но и для относительно небольших социальных и профессиональных слоёв населения, например, учёных.

Продолжая данную мысль, в рамках настоящей статьи опишем концепцию нового программного обеспечения, способного решить целый ряд социально значимых проблем в области российской, возможно, и мировой науки.

Прежде, чем приступить к описанию целей и задач, решаемых программным обеспечением, предлагаемым к рассмотрению в данной статье (далее «программой»), рассмотрим историю возникновения программы GitHub, функциональными возможностями которой изначально была вдохновлена концепция, описанная далее.

В 2007 г. группой программистов (Chris Wanstrath, P. J. Hyett, Tom Preston-Werner and Scott Chacon) в Сан-Франциско была разработана программа GitHub [34]. Она ориентирована на профессиональное сообщество программистов и решает следующий спектр задач:

1) Позволяет хранить интеллектуальные наработки (программного кода). Поскольку программа основана на облачных технологиях, пришедших в Россию по мере развития интернета, то часть интеллектуальных наработок программистов стала храниться в открытом доступе, представляя собой некий виртуальный сервер (место хранения информации);

2) Обеспечивает условия для совместной работы специалистов. Программа позволяет программистам писать код совместно, делясь им между собой посредством ссылки на репозиторий (гиперссылки в интернете). Это означает, что над одной и той же частью программы может работать одновременно несколько программистов, что делает процесс работы параллельным, значит, повышает количество одновременно задействованных специалистов, могущих работать над одной и той же программой;

3) Позволяет осуществлять ретроспективные перемещения по версиям. Программа позволяет программисту перемещаться не только к самой последней версии программного кода, но и просматривать и возвращаться к тому, над чем он работал, например, неделю, месяц или год назад. Таким образом, он может анализировать процесс эволюционирования программы в ретроспективе;

4) Обеспечивает публичный доступ к программному коду. Программный код, опубликованный на GitHub, в большинстве случаев сразу становится общедоступным, что позволяет знакомиться с ним и дорабатывать его своими силами любому желающему программисту (т. н. «открытый, исходный код»).

На февраль 2020 г. данной программой пользуется 40 млн. пользователей, а на самой платформе зарегистрировано 100 млн. репозиторий (виртуальных мест для хранения программного кода) [32].

Поскольку настоящая статья ориентирована преимущественно на читателей из гуманитарных областей наук (социологии, психологии, философии, педагогики и др.), то для увеличения корректности понимания описываемых идей было решено минимизировать объём используемых профессионализмов, состоящих из множества технических терминов и англоязычных аббревиатур, и описать именно концептуальную сторону программы.

Второй причиной, исходя из которой было принято решение акцентировать внимание на концептуальной стороне программы, служит то, что технические термины и синтаксис языков программирования – которыми нередко сопровождаются статьи, описывающие наработки в области программной инженерии, – являются искусственными языками [26, с. 81], созданными для упрощения общения человека с компьютером или с другими программистами, заставшими детерминированный по времени этап развития информационных технологий. Подобные языки как являющиеся достаточно молодыми эволюционируют на порядки быстрее естественных языков. Последнее представляет собой большой недостаток с той точки зрения, что синтаксис (совокупность терминов и символов, посредством которых программист сообщает программе требуемые команды) быстро устаревает, вследствие чего быстро бы устарела и сама статья, что усложнило бы её восприятие спустя несколько лет/десятилетий.

Наконец, третья причина акцентирования внимания на концептуальной стороне программы состоит в том, что одну и ту же функциональность программы возможно реализовать с помощью разных языков программирования, что делает описание функциональности и решаемых ею проблем первичнее деталей реализации.

ОСНОВНЫЕ ПОНЯТИЯ

Прежде чем перейти к основному содержанию настоящей статьи, договоримся о понятийном аппарате, который будет использоваться в дальнейшем.

Гениальное открытие – при употреблении данного термина мы будем опираться на определение, сформулированное профессором С.В. Черновым как «своеобразное *про-видение* (курсив С.В. Чернова. – *Р.М.*), это некий слепок, след, отпечаток, символ его (гения. – *Р.М.*) духовной жизни» [28, с. 79]. Причём – и это не менее важно – требуется наделять именем гения конкретных личностей лишь по прошествии длительных, исторических периодов времени, измеряемых столетиями и тысячелетиями.

Полагаем, что, как и множество др. статистически редких социальных явлений, к которому можно отнести явление гениальности, выборка индивидов, расположенных к совершению гениального открытия, в рамках генеральной совокупности всех пользователей программы будет составлять не более доли, выходящей за пределы $\pm 3\sigma$ [31] (т. е. лишь 0,26% от генеральной совокупности).

Однако, поскольку наша программа носит в том числе и социологический характер, то «История гениального открытия» – вполне подходящее название, указующее остальным пользователям в лице учёных, составляющим статистическое содержание первых 3σ , вектор интеллектуального развития, следование которому впоследствии рано или поздно приведёт их к заслуженному положению соответствующей социально-исторической роли, именуемой «гением», а их открытие станет «гениальным».

Программа (программное обеспечение, ПО, сайт, приложение, одностороннее веб-приложение или SPA) – это профессиональное сокращение от «программного обеспечения», термина, обозначающего информационную систему, состоящую из множества функций, объединение которых способствует решению тех или иных коммерческих, научных или др. задач в рамках конкретной *предметной области*, т. е. программа способствует достижению ставящихся целей за счёт использования информационных технологий (от англ. *Information Technology*, или сокр. *IT*).

В рамках настоящей статьи под этим понятием мы будем подразумевать преимущественно программу, рассчитанную на представителей научной сферы

деятельности, хотя функциональные возможности программы не будут исключать и возможности её адаптации для упрощения труда деятелей культуры, например, писателей, художников, музыкантов и т. д.

Пользователь (учёный; или «деятель культуры», или «гений», или «выдающийся индивид») – это лицо, взаимодействующее с программой. Поскольку изначально концепция программы «История гениального открытия» была ориентирована на научную сферу деятельности, то, чаще всего, под этим термином мы будем подразумевать учёного, т. е. человека, занятого научной деятельностью.

При этом планируется не отделять понятие *учёный* от понятий *преподаватель образовательного учреждения* и *научный сотрудник*. Под понятием *учёный* мы подразумеваем любого, кто связан с деятельностью, характерной для научной сферы, т. е. публикацией научных статей, написанием научных монографий и т. д.

Рукописи (труды, интеллектуальные наработки или личные дневники) – под данным понятием мы будем подразумевать всё, что в рамках использования программы было когда-либо создано, прочитано, обновлено или удалено её пользователем.

Согласно ГК РФ, ст. 1225, п. 1, интеллектуальные наработки также обозначают «интеллектуальной собственностью» [5], однако мы не будем использовать последнюю, правовую формулировку, поскольку твёрдо убеждены, что хотя деятельность представителей науки и культуры и является индивидуализированной и всегда имеет своего автора или коллектив таковых, однако уровень влияния рукописей данных индивидов на интеллектуальный прогресс социума носит массовый характер и является результатом реализации множества причинно-следственных связей, в которых участвовали и др. индивиды (например, меценаты, министры, директора НИИ и т. д.), цепь взаимодействий с которыми привела к гениальному открытию. По этой причине гениальное открытие есть явление отчасти групповое, имеющее множество неизвестных авторов.

Поскольку под понятием *рукопись* будут подразумеваться любые труды, когда-либо написанные учёным, в том числе могущие не иметь отношения к науке, но которые последний, тем не менее, согласился сохранить в программе,

то в это понятие мы также будем включать *личные дневники и частные письма* учёного.

ЗНАЧЕНИЕ ПРОГРАММЫ ДЛЯ НАУКИ, КУЛЬТУРЫ И ОБРАЗОВАНИЯ

Изначальная идея создания программы «История гениального открытия» [33] преследовала **три основные цели**:

1) создать условия для хранения интеллектуальных наработок пользователей на длительном (измеряемом несколькими десятилетиями/столетиями) временном интервале времени; иными словами, сохранить и уберечь от временной гибели научное наследие российских ученых;

2) создать платформу, в рамках которой будет производиться большое количество интеллектуальных наработок, совокупность которых со временем начнёт представлять интерес для учёных, занимающихся исследованиями в области т.н. «больших данных» (от англ. *Big Data Science*). Чем больше будет становиться массив рукописей, тем больший интерес они могут начать представлять для статистов и исследователей явлений творчества и гениальности, тем меньшая статистическая погрешность будет образовываться при заключении выводов на основе данных материалов [21, с. 11]. Это позволит приблизиться к выявлению закономерностей создания научных теорий посредством получения количественных характеристик:

- о средней продолжительности создания научной теории;
- о среднем количестве этапов (версий), через которое проходит теория, чтобы достичь финального оформления;

- о средней продолжительности этапов над каждым этапом проектирования научной теории; так, например, может оказаться, что наиболее существенным по времени этапом является этап сбора первичных данных, а наименее продолжительным – рецензирование последней готовой версии рукописи.

Большие данные, собираемые в том числе по каждому отдельному учёному (здесь речь идёт о т. н. «версионировании»: программа будет хранить не только последнюю версию рукописи, но и все предыдущие версии), позволят видеть в ретроспективе, как изменялось мировоззрение учёного в течение жизни, как на ход его мысли влияло общение с другими учёными. Для крупных учёных часто характерно общение с себеподобными, не менее крупными научными

деятелями. Так, например, С.П. Капица, создатель микротрона и популяризатор российской науки, в течение жизни общался с классиком российского авиационного А.Н. Туполевым, что вполне могло найти отражение в его научных работах.

3) предоставить автоматизированный инструментарий для упрощения интеллектуальной деятельности учёных. Иначе говоря, функциональность программы будет представлять собой набор функций, использование которых будет направлено на повышение производительности интеллектуального и творческого труда пользователя.

Достижение вышеописанных целей характеризует следующую ценность программы:

1) с точки зрения науки, программа увеличивает производительность труда учёных;

2) с точки зрения культуры, программа позволяет сохранить интеллектуальные наработки со временем, тем самым обеспечив передачу государственного, культурного наследия последующим поколениям;

3) с точки зрения образования, которое во многом базируется на результатах деятельности относительно небольшой группы общепризнанных, мировых учёных в разных областях наук, именуемых обычно т. н. «классиками», программа позволит сохранить многообразие научных трудов, незначительная часть которого впоследствии начнёт выступать в роли образовательных материалов для будущих, молодых поколений педагогов.

СТРАНИЦЫ (СМЫСЛОВЫЕ РАЗДЕЛЫ) ПРОГРАММЫ

Одним из способов, способствующих взгляду на программу с высокого уровня абстракции, является просмотр списка страниц программы (сайта) как смысловых разделов, обладающих некоторой завершённой функциональностью. Поэтому при описании функциональности программы далее будут приведены именно её страницы. Отметим, что для минимизации технической, служебной информации из таблицы были исключены такие разделы, как страницы авторизации, политики конфиденциальности, ознакомительной информации о программе и т. д. Итак, программа будет содержать следующую структуру (см. табл. 1).

Таблица 1. Страницы (разделы) программы «История гениального открытия»

№	Название страницы	Описание раздела
1	Регистрация	Это вторая страница, которую пользователь видит, знакомясь с данной программой. На странице регистрации у пользователя запрашивают следующие данные: ФИО, email, пароль, дата рождения (для отправки поздравлений и определения возраста), адрес проживания, ученая степень и ученое звание, поле исследовательских интересов.
2	Список рукописей	Здесь списком будут показаны все материалы, созданные пользователем: научные публикации, монографии, учебники и др., причем можно будет открыть доступ любому другому пользователю для совместной работы над материалом(-ами).
3	Добавить рукопись	Аналог классического редактора текста «Microsoft Word», но с рядом дополнений, характерных для исследовательских потребностей научного сообщества.
4	Обработка источников	На этой странице будет храниться набор полезных интернет-сервисов, упрощающих выполнение рутинных задач в научно-исследовательской деятельности. Их пользователь может использовать, например, для автоматизированной обработки собственных текстов и вторичных источников (проверки стиля, орфографии, контент-анализа и т. д.).

№	Название страницы	Описание раздела
5	Генератор аудио-монографий	На данной странице будет находиться сервис, позволяющий перекодировать электронные тексты в формат аудиозаписей. До появления концепции программы было замечено, что большинство монографий не имеет голосовых версий, вследствие чего учёные теряют возможность экономии научно-исследовательского времени, приобретаемой посредством распараллеливания деятельности.
6	Наследники рукописей	Поскольку для ряда крупных учёных, разрабатывающих теории парадигмального уровня, характерна самостоятельная работа над научными открытиями, то данная страница позволит им спланировать дальнейшую судьбу совершенных ими опытов и открытий в случае преждевременного ухода из жизни. На данной странице пользователь может приложить оригинальную версию завещания с подробным описанием лиц, имеющих, на его взгляд, заслуженное отношение к открытию.
7	Научное становление	Здесь, по желанию учёного, планируется фиксация интернет-сайтов, посещаемых последним. Это позволит составлять анонимный рейтинг наиболее популярных порталов среди учёных, а также сохранять историю перемещений в интернете, отражающую научное становление последнего.
8	Научная переписка	Социальная сеть для общения ученых и создания онлайн-форумов, направленных на обсуждение научных проблем и поиска оптимальных путей их решения.

№	Название страницы	Описание раздела
9	Рукописи коллег	Данная страница пока находится под вопросом, поскольку на ней с разрешения авторов изначально планируется опубликовать рукописи других учёных. При её проектировании мы опирались на работы инженера и автора теории решения изобретательских задач (ТРИЗ) Г.С. Альтшуллера, согласно которым средняя задержка между созданием научно-технической идеи и её промышленной реализацией составляет около 10 лет [1]. Страница была призвана сократить эту задержку посредством открытого доступа к черновикам учёных, однако это противоречит закону об авторском праве.
10	Научный самоанализ	Здесь учёный сможет просматривать статистику по своей научной-творческой продуктивности. Например: как часто он возвращается к рукописям, какие рукописи важны для него больше всего, сколько времени, в среднем, он посвящает работе над одной рукописью и т. д.
11	Контакты коллег	Поскольку в научном мире на 2019 г., чаще всего, по сложным, научным вопросам принято общаться через email-почту, то одна из страниц сайта будет содержать список email-почт российских и зарубежных ученых с их разделением по научным направлениям, ученой степени и ученому званию. Данный список может формироваться по одному из алгоритмов на выбор: либо с согласия самих пользователей программы, либо посредством написания дополнительной программы для сбора контактов ученых, открыто размещенных на сайтах вузов в интернете.

№	Название страницы	Описание раздела
12	Научный коллектив	Здесь будет представлена функциональность, позволяющая сразу нескольким учёным работать над одной рукописью одновременно, что является одной из ключевых функций GitHub.
13	Личный дневник	Здесь учёный может фиксировать мысли, накопившиеся в течение дня, недели или других интервалов времени, если считает, что их можно доверить только личному дневнику. По умолчанию, все дневники должны храниться как информация, скрытая от других пользователей, а сам текст должен шифроваться по специальному алгоритму, который пока находится на стадии разработки.
14	Мультимедийные факты	На этой странице пользователь может загружать в базу данных программы все, что позволяет персонализировать его личность, идентифицировать ее с созданными им творениями.

Поскольку разработка данной программы в настоящее время уже идёт полным ходом, то имеется некоторый *артефакт* в виде *интерфейса программы*, именуемого в профессиональной среде программистов как GUI (от англ. *Graphical User Interface*), позволяющий увидеть (см. рис. 1):

- как программа выглядит на момент написания настоящей статьи;
- какие смысловые разделы, показанные в левой части изображения, она содержит.

ФУНКЦИОНАЛЬНОСТЬ ПРОГРАММЫ

Часть интерфейса программы, ориентированного на работу с текстом, во многом будет напоминать классический редактор текста, например, Microsoft Word, но с рядом улучшений, актуальных для научной предметной области. Программа позволит хранить версии научных работ пользователя в облачном хра-

нилище интернета, к которому у пользователя будет иметься доступ с любого электронного устройства.

Иначе говоря, «История гениального открытия» будет предоставлять такую информационную среду, которая позволит хранить все интеллектуальные наработки ученых, а также работать совместно над одной рукописью, редактируя разные её части. Именно таким образом ведётся работа в коллективах программистов, когда каждый программист фокусируется на определённой части программы (точнее, на одном или нескольких модулях программного кода). Это может быть полезно учёным в процессе работы над коллективными монографиями или коллективной статьёй, где каждый пользователь отвечает за определённый смысловой блок, параграф или главу текста.

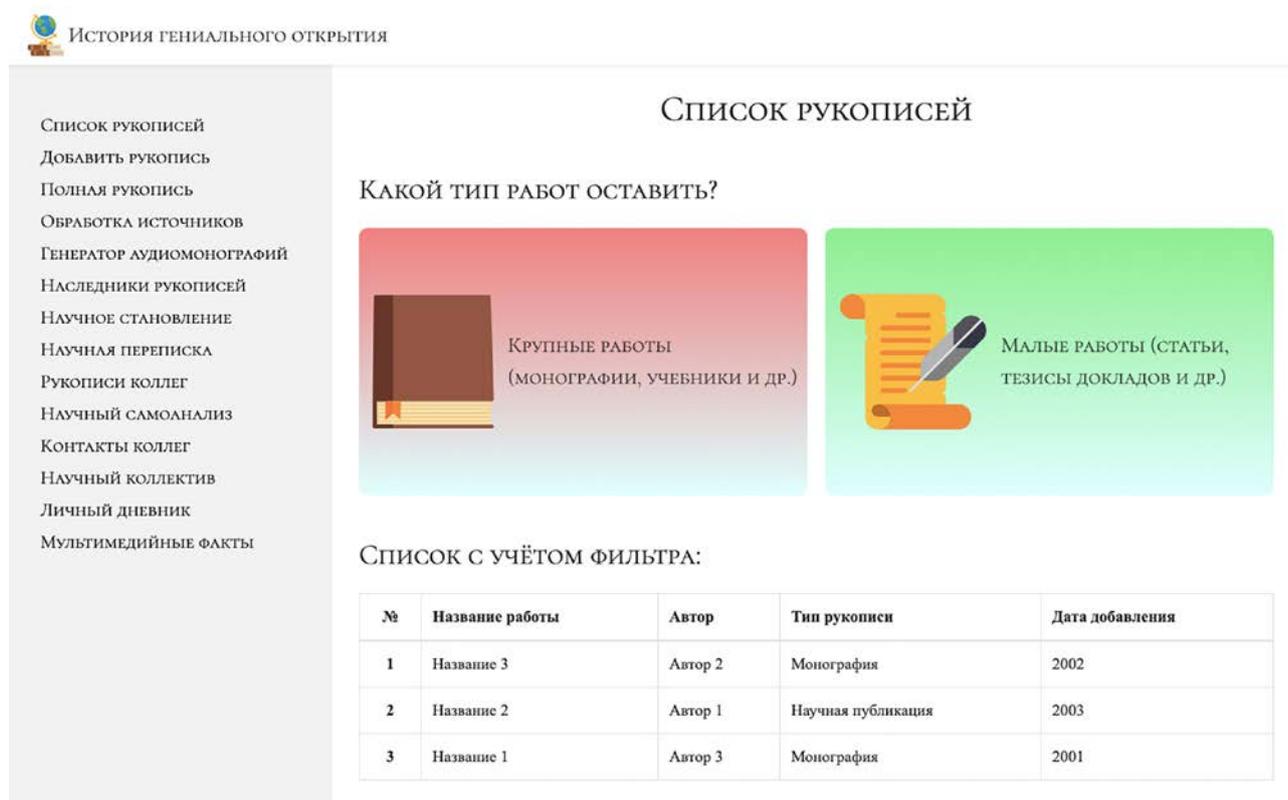


Рисунок 1. Текущий интерфейс (GUI) программы «История гениального открытия» на май 2020 г.

Справедливо будет задать следующий вопрос: «А зачем создавать новую программу для хранения информации, когда есть GitHub, на котором, по факту, уже можно хранить любой вид текстовой информации, в том числе и научные

работы?». Дело в том, что основной целевой аудиторией программы GitHub являются программисты, что предполагает концентрацию вокруг неё индивидов преимущественно технического склада ума, имеющих особые неформальную культуру, ценности и терминологию, нередко отличные от научных. Поэтому в течение года (апрель 2019 – май 2020) зарождалась идея создания отдельной программы для учёных – «История гениального открытия».

В конечном итоге программа должна будет представлять собой систему научно-творческого облачного хранилища, которая будет минимизировать вероятность потери важнейших научных открытий и новых разработок по ряду следующих возможных причин:

А. Преждевременный уход пользователя из жизни. Как таковой пример преждевременного ухода учёного из жизни, формализованный в биографической литературе, привести сложно, но мы можем представить несколько примеров из жизни деятелей культуры. Так, Л. да Винчи завещал часть художественных произведений своему ученику, Ф. Мельци [28, с. 528], имея, по ряду мнений, незавершённый портрет Джоконды, а Л. ван Бетховен – нотные партитуры своему племяннику К. ван Бетховену [9, с. 493], так и не успев завершить некоторые своих музыкальные работы.

Хотя нам не известен конечный результат посмертной передачи произведений вышеупомянутых деятелей культуры их близкому кругу лиц, не известны последствия этой передачи, однако уже из базовых постулатов социологии следует, что общество есть динамическая система, отличающаяся от кибернетических систем вероятностным характером осуществления социальных процессов [21, с. 7]. Значит, доверить рукописи компьютерной системе, при наличии надлежащих условий их хранения, достигаемых посредством резервного копирования данных, – более надёжное решение, нежели полагаться на изменчивый от природы т. н. «человеческий фактор». Тем более, что далеко не каждый наследник научных или культурных произведений способен грамотно оценить всю сложность и значимость переданных ему произведений, особенно если не имеет соответствующей тому образовательной и профессиональной подготовки.

Убеждены, что в ряде случаев «человеческий фактор» действительно играет важную роль и что ряду рукописей лучше было бы не «выходить в свет» в силу возможной:

- недоработанности идей, сосредоточенных в них;
- профессиональной халатности, могущей иметь место при их формулировании;
- малой достоверности оснований, на которых базируются научные утверждения;
- наконец, неготовности общества к принятию определённых научных концепций в силу недостаточного уровня развития цивилизации (пример: водородная бомба акад. А.Д. Сахарова и её использование в военном деле с отрицательными последствиями для общества).

Однако без создания подобной кибернетической системы учёные теряют возможность систематического исследования тех теорий, которые не «вышли в свет», теряется возможность исследования причин, по которым наследуемые рукописи оказались недостойными быть представленными в научной среде и обществе;

В. Особенности темперамента пользователя, вследствие которых научно-творческие разработки уничтожаются в короткий промежуток времени (пример: Н.В. Гоголь и его придание огню 2-го тома «Мертвых душ») [12];

С. Преждевременная поломка компьютера, отдельных его структурных элементов и невозможность вследствие этого восстановить хранимые на нём рукописи.

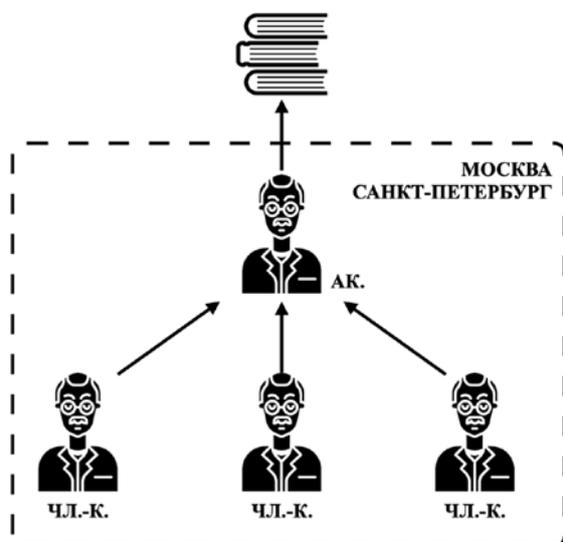
УСЛОВИЯ ВКЛЮЧЕНИЯ МЕХАНИЗМА ПОСМЕРТНОГО ОБНАРОДОВАНИЯ ДАННЫХ

Поскольку одна из основных целей программы будет состоять в том, чтобы сохранить рукописи пользователей и передать их последующим поколениям, то для этого должен быть разработан соответствующий алгоритм, который будет выступать связующим звеном между учёным и общественностью. Для этого был введён *алгоритм посмертного обнародования рукописей*. Иными словами, при уходе учёного из жизни его рукописи, будучи централизованно собранными в одну электронную папку (например, ZIP-архив), будут отправляться соответствующим

ющей группе специалистов, способных в полной мере оценить уровень вклада пользователя в российские науку или культуру.

Изначально планировалось, что таким связующим звеном должна быть организация, обладающая высшим научным авторитетом в сообществе учёных, где показатель авторитетности должен оцениваться количественными характеристиками данной социальной структуры (например, импакт-фактором научного журнала), а также состоящих в ней учёных (например, индекс Х. Хирша). Исходя из этого, по первоначальному замыслу более всего под данные параметры подходила Российская академия наук (РАН) [18].

ПРИМЕР ИЕРАРХИЧЕСКОЙ НАУЧНОЙ СТРУКТУРЫ, ОТВЕТСТВЕННОЙ ЗА ОЦЕНИВАНИЕ ЗНАЧИМОСТИ РУКОПИСЕЙ



ПРИМЕР РАСПРЕДЕЛЁННОЙ НАУЧНОЙ СТРУКТУРЫ, ОТВЕТСТВЕННОЙ ЗА ОЦЕНИВАНИЕ ЗНАЧИМОСТИ РУКОПИСЕЙ

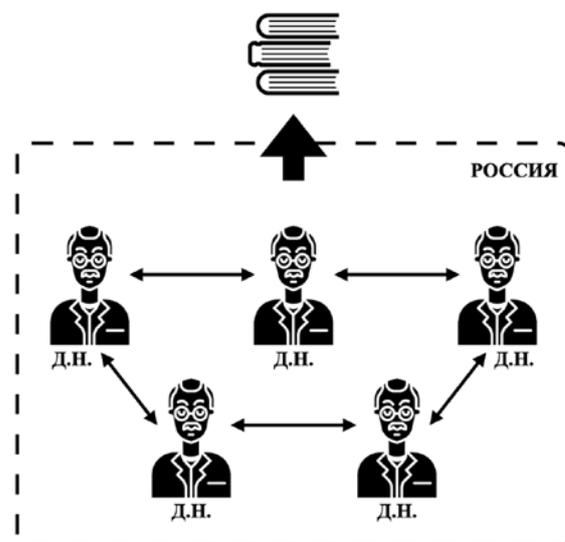


Рисунок 2. Схемы иерархической и распределённой структур оценивания рукописей.

Однако здесь следует заметить, что данный выбор пока не является окончательным, а подобная структура не обязательно должна быть централизованной и имеющей жёсткую структуру. Вполне возможно, что система распределённого рецензирования, где ресурсы и властные полномочия не сосредоточены в одних руках, а распределены между разными авторитетными деятелями науки (см. рис. 2) зарекомендует себя как более эффективная. *Главное основание: рецензирование посмертно обнародованных рукописей должны*

осуществлять специалисты соответствующих научного уровня и объёма знаний, причём таких специалистов должно быть несколько. Только таким образом можно будет достичь максимально справедливой оценки научного вклада, внесённого тем или иным пользователем на благо российской науки, и обеспечить все необходимые условия, чтобы рукописи незаслуженно «не канули в Лету».

У предлагаемой схемы распределённой редакционной коллегии, безусловно, есть ряд недостатков:

– например, у учёных из регионов, как правило, имеется меньший доступ к информации и знаниям, поскольку основной научный мейнстрим (от англ. *Mainstream*) во множестве крупных стран мира сосредотачивается именно в небольшом числе научных центров;

– также у учёных из регионов имеется меньше возможностей для участия в международных научных конференциях и изучения английского как языка, позволяющего коммуницировать с коллегами из других стран для обмена опытом;

– помимо прочего, не стоит забывать, что идеи для зарождения нового научного мейнстрима, исходя из «Структуры научных революций» Т. Куна, как правило, образуются в умах вполне конкретных учёных и становятся опубликованными в печатных или электронных изданиях лишь по прошествии, в лучшем случае, нескольких лет. «Тем не менее я и дальше буду говорить даже об отдельных открытиях (выделено нами. – Р.М.) как о революционных, поскольку только таким образом можно сравнить их структуру с характером, скажем, коперниканской революции, что и делает, по моему мнению, это расширенное понятие важным» [10, с. 31]. Последнее свидетельствует о важности географической близости к подобным учёным для своевременных поддержки и развития их научных идей.

Несмотря на вышеперечисленные недостатки, распределённая система оценивания рукописей всё же позволяет значительно минимизировать концентрацию научного авторитета в одном месте. Как решение можно внедрить систему распределения голосов, учитывающих все уровни учёных. Например, 50% голосов для решения дальнейшей судьбы рукописей будет вноситься представителями крупнейших в России научных центров, а другие 50% голосов – регио-

нальными научными субъектами (как специализированными институтами, так и отдельными научными авторитетами в виде конкретных лиц из регионов).

Итак, мы пришли к тому, что в случае преждевременного ухода пользователя из жизни все версии рукописей должны сжиматься в один, единый ZIP-архив и отправляться на электронную почту соответствующей научной структуры с темой:

Научное наследие России. Наука: S. Автор №: ID. ФИО автора: N.

где

S – от англ. *Science*, это название научного направления(-ий), в рамках которого(-ых) преимущественно трудился пользователь, например, биология;

ID – от англ. *Identification*, это автоматически генерируемый, уникальный номер автора относительно других авторов, который будет храниться в базе данных программы «История гениального открытия»;

N – от англ. *Name*, это фамилия, имя и отчество автора.

Также возможна отправка рукописей одновременно нескольким получателям, например, нескольким российским университетам или отдельным их институтам. Последнее, с одной стороны, повысит вероятность признания научных заслуг ушедшего из жизни учёного, с другой стороны, диверсифицирует риски [29], при которых его работа будет безвозвратно потеряна вследствие её хранения в одном месте.

Программа будет определять, что пользователь преждевременно ушёл из жизни на основании следующих факторов:

1) пользователь в течение 3 (трёх) полных лет не появлялся (авторизовался) в программе;

2) по прошествии 3 лет пользователь никак не отреагировал (например, посредством авторизации в программе) на 3 email-уведомления и 3 sms-сообщения, оповещающих о временной «заморозке» его аккаунта, каждое из которых отправлялось ему с перерывами в 1 года (первое – по итогам 1 года отсутствия в программе; второе – по итогам 2 лет; по итогам 3 лет).

Следует отметить, что вышеописанный алгоритм посмертного обнаружения данных в настоящий момент не совершенен и не учитывает простейших факторов смены или поломки технических устройств (смартфонов, ноутбуков и

др.), смены сим-карт, создания новых электронных почт и т. д. В идеале в алгоритм следует добавить также оповещение единомышленников пользователя, которых последний будет определять самостоятельно и которые, с высокой долей вероятности, имеют регулярную коммуникацию с ним (контактируют как минимум 1 раз в год посредством электронных писем или личного общения).

В случае с рукописями творческих деятелей, в частности, писателей, их работы могут отправляться в Российский союз писателей [19] (РСП) или любые другие творческие учреждения, расположенные на высшем иерархическом уровне определённого творческого сообщества.

НАУЧНЫЙ САМОАНАЛИЗ

Требования к осуществлению научной деятельности имеют свойство изменяться со временем. Одним из таких требований в эпоху информатизации (примерно с 1995 г. в России) стало требование уметь использовать на *базовом уровне* компьютерные технологии. Под базовым уровнем здесь, как правило, понимаются навыки работы с программами из пакета Microsoft Office (Word, Excel и др.) и браузерами, позволяющими выходить в интернет.

Когда учёный регулярно выходит в интернет, то его браузер сохраняет в своей истории записи о тех сайтах, которые тот посещал в течение определенного промежутка времени (например, в течение недели, месяца или года). Браузер делает это при условии, что учёным не изменялись его стандартные настройки.

На самостоятельном примере было обнаружено, что при регулярном выходе в интернет браузерная история по итогам года начинает хранить базу в 62 050 запросов⁹ (посещённых страниц сайтов). Данное число формируется при среднем показателе в 170 запросов в день, помноженном на 365 дней в году.

Иначе говоря, по прошествии одного года пользования интернетом браузер сохранил множество сведений о поисковой активности, объём которых достаточен, чтобы именовать данный массив запросов большими данными. Это

⁹ Прим. авт.: автор на момент написания статьи обучается по специальности разработчика программно-информационных систем, что предполагает регулярную работу с информационными технологиями и чтение документаций в интернете. Данный фактор существенно влияет на количество совершаемых интернет-запросов.

значит, что, сохранив массив данных активностей в отдельный файл, можно работать с ним как с *репрезентативным источником первичных данных* об интересах учёного и типах источников, к которым последний обращается в процессе проведения научных исследований.

Например, в отношении данного массива можно применить т. н. контент-анализ, являющийся одним из видов «количественных методов анализа документов» [17, с. 292] в социологии. После применения такого анализа получим список наиболее часто набираемых запросов (или ключевых слов), количество их набора, а также список сайтов, посещаемых пользователем. Далее возможна типологизация полученных ключевых слов и сайтов.

Поскольку данный вид анализа – контент-анализ сайтов, посещённых в интернете, – насколько нам известно, ранее никем не использовался и поскольку этот вид анализа отличается от классического контент-анализа, предполагающего работу с бумажными или электронными документами, то мы вправе выделить его в отдельный вид анализа, берущий своё начало в эпохе информатизации, ознаменовав его как *браузерный контент-анализ*.

Стоит заметить, что подобная функция должна внедряться только после тщательного изучения нормативно-правовых документов, касающихся персональных данных, например, документа от 27.07.2006 № 152-ФЗ «О персональных данных» [25], поскольку история посещения сайтов пользователем относится к категории информации частного характера.

Также пока не был изучен технически вопрос о возможности интеграции данной функции с самими браузерами (Яндекс.Браузером, Google Chrome, Mozilla Firefox и др.):

- согласятся ли браузеры предоставлять возможность экспорта данных пользователя;
- если согласятся, то на каких условиях;
- какие данные среди посещённых сайтов пользователя программа должна игнорировать (фильтровать), исключая их из анализа. Например, корректно ли будет, если программа станет хранить сайты, касающиеся медицинских деталей жизни пользователя (например, запросы, касающиеся его заболеваний или заболеваний членов его семьи). С точки зрения медицины, это некорректно. Ещё

в клятве Гиппократата, содержащей основные положения медицинской этики, говорится: «Что бы при лечении – а также и без лечения – я ни увидел или ни услышал касательно жизни людской из того, что не следует когда-либо разглашать, я умолчу о том, считая подобные вещи тайной» [3, с. 88];

– если согласие не будет получено, то какой должна быть инструкция, объясняющая пользователю, как выгружать и осуществлять браузерный контент-анализ самостоятельно.

Все вышеперечисленные вопросы требуют отдельного изучения.

АНАЛИЗ ИССЛЕДОВАТЕЛЬСКИХ ИНТЕРЕСОВ УЧЁНЫХ

Другая разновидность контент-анализа, а именно, *монографический контент-анализ*, может пригодиться ученым в другой ситуации. В науке, хотя и в неявном виде, существует проблема всестороннего ознакомления с идеями других учёных, быстрого «вхождения в курс дела». Она особенно актуальна на этапах формирования базы источников для проведения исследования, а также при обработке вторичных данных.

Причины, по которым учёному может потребоваться быстрое, всестороннее ознакомление с идеями других авторов:

- они не работают(-ли) вместе на протяжении длительного времени;
- ранее не изучались письменные труды другого автора (учебные пособия, монографии, научные статьи и т. п.);
- *поле исследовательских интересов* конкретного учёного не указано в явном виде на его профессиональной странице (например, на сайте университета, в котором тот преподает).

Все факторы, перечисленные выше, ведут к тому, что, чтобы понять основные положения научной точки зрения другого автора, учёному требуется выделить время, как правило, измеряемое несколькими днями/неделями, а затем прочесть ключевые работы интересующего учёного. А если учесть, что при написании одной научной статьи учёный в среднем упоминает около 15–20 фамилий в списке используемых источников, то последнее увеличивает время, необходимое для предварительного ознакомления с рукописями других авторов в несколько раз. Поэтому для минимизации времени, затрачиваемого на предварительное изучение работ других авторов, предлагается использовать специаль-

ный микросервис внутри программы, который будет брать список рукописей конкретного автора и проводить по каждой из рукописей контент-анализ. Последнее позволит автору значительно быстрее принимать решение о том, стоит ли углублять изучение идей конкретного автора, схожа ли его точка зрения с точкой зрения пользователя.

Рассмотрим следующий пример. Нами был проведен контент-анализ монографии «Книга о гениальности. Т. 1: Человеческий гений: Природа. Сущность. Становление. Монография» [28] профессора С.В. Чернова с помощью свободно распространяемой зарубежной программы для контент-анализа «Voyant Tools» [16], использование которой (либо ее улучшенного, дополненного аналога) планируется встроить в интерфейс программы «История гениального открытия».

Таблица 2. Результат контент-анализа монографии С.В. Чернова «Книга о гениальности. Т. 1: Человеческий гений: Природа. Сущность. Становление. Монография»

№	Слово	Упоминаний, упом./мон.	Слово	Упоминаний, упом./мон.
1/11	человека	510	всегда	218
2/12	гения	456	Леонардо	200
3/13	гений	427	Бальзак	193
4/14	человек	368	более	188
5/15	жизни	296	путь	181
6/16	лишь	293	здесь	177
7/17	людей	287	человеческого	174
8/18	Бердяев	270	даже	162
9/19	творчества	239	можно	161
10/20	жизнь	223	творчество	161

Видим (см. табл. 2), что наиболее часто встречающимися понятиями в работе профессора С.В. Чернова являются понятие о человеке («человека» – 510 сл./мон.; «человек» – 368 сл./мон., «жизни» – 296 сл./мон., «творчества» – 239 сл./мон.; *сл./мон.* – это введённое автором сокращение, обозначающее меру измерения ключевых слов монографии, расшифровывается как: слов/словосочетаний в монографии). Иначе говоря, как кандидат педагогических наук и доктор философских наук, С.В. Чернов смотрит на явление гениальности с гуманитарной точки зрения, ставя во главу угла человека, его жизнь и результаты творчества, составляющие культурное наследие общества.

Также монографический контент-анализ позволяет установить, что большое влияние на формирование научной точки зрения С.В. Чернова по проблематике гениальности оказали труды философа Н.А. Бердяева («Бердяев» – 270 сл./мон.). Безусловно, последнее можно было бы установить и при беглом анализе списка используемых источников, приводимого обычно в конце работ. Однако, во-первых, это потребовало бы ручного подсчёта, во-вторых, здесь не учитывается фактор, что в списке используемых источников были приведены лишь 1–2 работы конкретного автора, являющиеся ключевыми. Истории науки известно множество случаев, когда учёный после себя оставлял всего одну крупную работу перед уходом из жизни. Последнее особенно характерно для учёных, ушедших из жизни в молодом возрасте.

Другая, противоположная точка зрения на проблематику гениальности содержится в монографии «Изменчивость и гениальность» профессора, доктора биологических наук С.В. Савельева (см. табл. 3) [15]. Противоположна она как минимум в том, что С.В. Савельев рассматривает явление гениальности с точки зрения устройства головного мозга (нейрофизиологии) и поведения индивида (бихевиоризма), о чем свидетельствует результат контент-анализа его работы («мозга» – 810 сл./мон.; «области [мозга]» – 177 сл./мон.; «полей» – 123 сл./мон.; «поведения» – 93 сл./мон.). На долю «человека», как видим, выпало всего 304 упоминания, что в 2,66 раз меньше, чем количество упоминаний о мозге. Личность и ее так называемый *богатый внутренний мир* для профессора С.В. Савельева, исходя из результата анализа, второстепенны и мало влияют на

гениальное становление отдельно взятого индивида. Гениальность, с данной точки зрения, есть механизм, случайно predetermined природой.

При использовании контент-анализа важно «держать в уме» объект и предмет, которым посвящена монография, обычно выражаемые в названии работы. Так, например, если бы в рамках данной статьи анализировалась не монография С.В. Чернова «Книга о гениальности. Т. 1: Человеческий гений: Природа. Сущность. Становление. Монография», а, например, другая его работа, «История гениальности в России. Книга первая: Университет всея Руси. Монография» [27] – посвященная великому русскому учёному М.В. Ломоносову, и анализу основных факторов, поспособствовавших его гениальному, научно-творческому становлению, – то, видя преобладающую долю упоминаний М.В. Ломоносова в данной работе, справедливо было бы сказать о т. н. «смещении выборочной совокупности» (единиц анализа) [17, с. 204], поскольку данная работа была изначально посвящена упомянутому объекту анализа (М.В. Ломоносову).

Таблица 3. Результат контент-анализа монографии С.В. Савельева
«Изменчивость и гениальность»

№	Слово	Упоминаний, упом./мон.	Слово	Упоминаний, упом./мон.
1/11	мозга	810	очень	82
2/12	человека	304	неокортекса	80
3/13	области	177	поверхности	80
4/14	мозг	168	системы	79
5/15	полей	123	мозге	75
6/16	людей	120	является	75
7/17	поля	112	головного	74
8/18	поведения	93	гениальности	70
9/19	коры	84	более	69

№	Слово	Упоминаний, упом./мон.	Слово	Упоминаний, упом./мон.
10/20	между	84	изменчивости	67

Имея в руках такой инструмент, как монографический контент-анализ, исследователь может относительно быстро отбирать необходимые для анализа вторичные источники и быстро фильтровать их на предмет схожести/различий точек зрения. Данный вид контент-анализа также может упростить процесс подбора ключевых слов к научным публикациям, возможно, и вовсе полностью исключить необходимость в их подборе при полной автоматизации процесса.

Перед началом разработки данного микросервиса следует обратить внимание на авторское право, предварительно определив:

– какие именно работы в программе могут храниться по умолчанию (например, считающиеся т. н. «общественным достоянием», загруженные после 70 лет с момента ухода учёного из жизни [6]);

– какие работы пользователь сможет загружать для анализа самостоятельно, например, при наличии личной договорённости с автором рукописи.

ДОБАВЛЕНИЕ ЕДИНОМЫШЛЕННИКОВ В СПИСОК ПОСМЕРТНЫХ ПОЛУЧАТЕЛЕЙ РУКОПИСЕЙ

Для жизни гениев обычно характерно наличие избранного круга единомышленников, поддерживающих их идеи при жизни. Как правило, данные индивиды выступают в роли учеников или меценатов гениев, перенимают их знания, опыт и идеи.

Наличие единомышленников возможно проследить на множестве примеров:

– в жизни «отца теоретической космонавтики» К.Э. Циолковского единомышленником был А.Л. Чижевский [8, с. 108];

– у изобретателя М.В. Ломоносова – И.И. Шувалов [18, с. 59];

– у художника Л. да Винчи – Ф. Мельци [28, с. 528];

– у философа Сократа – Платон [20, с. 8];

– у одного из основоположников Computer Science Дж. фон Неймана – А. Тьюринг [22, с. 35-37];

– у музыканта П.И. Чайковского – Н.Ф. фон Мекк [14, с. 77];

– у физика, создателя микротрона и популяризатора советской науки С.П. Капицы – А.Н. Туполев [13].

Наблюдая за участью множества рукописей выдающихся индивидов, доктор исторических наук, профессор И.Н. Данилевский пишет: «Что касается дневников как автокоммуникативных источников, частной переписки, имеющей конкретного адресата, или мемуаров-автобиографий, нередко адресованных прямым потомкам, то они зачастую если и не уничтожались "самим автором" или адресатом, то хранились небрежно и гибли, и если даже бережно сохранялись потомками, то системы их приема на государственное хранение (в отличие от, например, делопроизводственных источников, также не предназначавшихся для печати) не существовало, да и не существует до сих пор. Попасты в архив они могли, как правило, в составе личных фондов, а еще чаще – коллекций, что затрудняет их поиск и использование» [7, с. 467].

По этой причине в программе должна быть предусмотрена возможность создания списка единомышленников, каждому из которых она автоматически будет отправлять рукописи пользователя после его ухода из жизни. Это позволит минимизировать вероятность бесследного исчезновения идей, потенциально важных для российской научной истории. Последнее будет достигаться за счёт *диверсификации рисков* (понятие, активно используемое в области финансовых инвестиций [23]), т. е. передачи рукописей дальше в соответствии с теми социальными связями, которые установит сам пользователь, либо тех, кому будет предоставляться право на это (соответствующая роль в программе).

Следует отметить, что наличие списка единомышленников не противоречит праву на наследование, установленному государством [4], оно лишь дополняет его. При отправке официального запроса с документом, подтверждающим право наследования, члены семьи должны иметь возможность получения рукописей на равных с единомышленниками, установленными пользователем программы.

ДОПОЛНИТЕЛЬНАЯ ФУНКЦИОНАЛЬНОСТЬ

Далее будут описаны некоторые другие микросервисы программы, которые не были подробно раскрыты ранее и на которые стоит обратить внимание.

1. Хранение истории интернет-перемещений учёного на различных сайтах как способ фиксации научного становления. Ранее данный пункт уже раскрывался, когда рассматривалась полезность браузерного контент-анализа как способа научного самоанализа. Хранение в программе первичных данных с историей посещённых сайтов в будущем, после ухода учёного из жизни, может представлять собой совокупность фактов о научном становлении последнего.

2. Распознавание скан-копий текстов в формате DJVU и перевод их в текстовые, редактируемые форматы (TXT, DOCX). Здесь речь идёт о создании такого микросервиса в программе, который будет способен распознавать и конвертировать тексты монографий в формат, удобный для работы над рукописью. Монографии в формате DJVU, как правило, представляют собой работы устаревшего формата, которые были изданы примерно до 1990 г., а затем переведены в электронный вид посредством сканирования. Данный микросервис будет решать проблему ручного перепечатывания текста, обычно актуальную в ситуациях, когда учёному требуется сослаться на фрагмент вторичного источника.

3. Возможность ведения личных дневников. Здесь учёный сможет фиксировать мысли, накопившиеся в течение дня, недели, месяца, года или любых других интервалов времени, если посчитает, что их можно доверить только личному дневнику. Ряд биографий выдающихся индивидов показывает, что ведение дневников – довольно характерная черта для выдающихся индивидов, сами результаты которой (дневники) могут представлять впоследствии интерес для биографов.

По умолчанию, все дневники должны храниться в программе как частная, закрытая информация, и пользователь не должен иметь возможности сделать её публичной. Для разработки последнего микросервиса также потребуется задействовать современные достижения *криптографии*, позволяющие зашифровывать и расшифровывать текст [30, с. 15]. Подобно тому, как великий художник эпохи Возрождения Л. да Винчи зашифровывал страницы личных дневников, которые можно было прочесть «лучше всего с помощью зеркала» [28, с. 338], тем

самым минимизируя риск их прочтения третьими лицами, так же и алгоритмы криптографии будут призваны защищать персональные данные от посторонних глаз.

Однако видится менее целесообразным и даже рискованным использовать шифрование в отношении к основным рукописям пользователя, его научным работам. Риск заключается, прежде всего, в том, что в случае утери алгоритма дешифровки текста (определённого модуля программного кода, отвечающего за перевод текста на человеческий язык) есть вероятность безвозвратной утери рукописей сразу всех пользователей, поскольку без специального ключа-дешифратора рукописи будут представлять собой не более, чем набор непонятных комбинаций символов. Помимо прочего, добавим, что научные работы, как правило, изначально ориентированы на публичный просмотр, а также, учитывая специфику предметной области (научная деятельность), устаревают очень быстро, значит, не имеет значения и само шифрование.

Относительно скорости устаревания идей в науке достаточно точно подметил основоположник «понимающей социологии», немецкий профессор М. Вебер в 1918 г. в своей речи в Мюнхенском университете: «Напротив, каждый из нас (учёных. – *P.M.*) знает, что сделанное им [открытие] в области науки устареет через 10, 20, 40 лет. Такова судьба, более того, таков смысл научной работы, которому она подчинена и которому служит, и это как раз составляет её специфическое отличие от всех остальных элементов культуры; всякое совершенное исполнение замысла в науке означает новые "вопросы", оно по своему существу желает быть превзойденным» [2, с. 532].

4. Синхронизация с электронной почтой и хранение частных переписок для биографов. В рамках данного микросервиса будет накапливаться история частных писем учёного, которые в дальнейшем смогут изучить исследователи его творчества в процессе написания биографической монографии. Для реализации данной функции потребуются техническая интеграция с такими почтовыми сервисами, как Яндекс и Gmail. Особенно необходима интеграция с Mail.ru, почты на данном сервисе имеет целое поколение профессоров.

На данный момент автору неизвестно, как реализовать подобную интеграцию и согласятся ли перечисленные выше компании на неё, поскольку здесь

затрагивается вопрос личных данных пользователей (учёных), вследствие чего внедрение данного микросервиса требует продолжительной подготовительной работы. Возможно, что более удачным решением, требующим минимума интеграций, станет разработка собственного микросервиса в пределах программы, позволяющего отправлять и принимать электронные письма.

5. Научный самоанализ. В этом разделе учёный может просматривать сводную статистику по своей научной эффективности. Например, посмотреть то, как часто он возвращается к рукописям, какие рукописи актуальны для него больше всего, сколько времени он, в среднем, работает над одной рукописью.

6. Небольшие сервисы для обработки рукописей и вторичных источников. Поскольку большим преимуществом использования информационных технологий является то, что они позволяют экономить время на автоматизации рутинных процессов, то программа «История гениального открытия» будет снабжена рядом гиперссылок на дополнительные внешние сервисы, призванные сократить временные расходы исследователей на ручных операциях.

В программе планируется реализовать следующие функции:

- перекодировка изображений текстов (скриншотов, фотографий) в текстовый формат;
- удаление разрывов между строками, т. е. удаление всех лишних переносов в скопированных частях PDF-файлов;
- разделение PDF-файлов на части, позволяющие отправлять студентам только те части (страницы) монографий, которые нужно изучить к следующим занятиям;
- обработка текста для его более качественного, презентабельного вида перед отправкой в научный журнал или печать;
- ряд других сервисов, которые будут особенно полезны представителям гуманитарных наук, позволяющих анализировать социальные проявления индивидов в интернете.

ЭТИКО-ПРАВОВЫЕ АСПЕКТЫ ПРОГРАММЫ

Ещё один вопрос, который требует тщательного предварительного изучения перед началом разработки данной программы, – это вопрос *научной этики*. Например, одна из функциональных возможностей программы предполагает,

что учёный сможет хранить личные дневниковые записи в электронном виде при использовании ПО. В связи с этим становится актуальным вопрос границ, за пределами которых мысли учёного перестают быть т. н. «достоянием общечеловеческой культуры».

Чтобы отобразить основные различия в уровнях доступа к информации учёного, введём три основных класса данных, с точки зрения обеспечения их конфиденциальности:

1) открытые данные. Доступ к ним может иметь, в зависимости от настроек системы, либо любой человек, который зашёл на сайт, либо только те его посетители, кто зарегистрированы на нём (большинство зарегистрированных пользователей, по изначальному замыслу, будет представлять собой учёных). Под этим классом данных планируем хранить завершённые либо черновые рукописи учёных, на которые последние дали согласие для публикации на сайте;

2) полузакрытые данные. Данный класс предполагает, что доступ к таким данным не будет иметься у внешних пользователей программы (учёных), но будет у администраторов базы данных (АБД). К этому классу данных планируем относить, например, черновые рукописи учёных;

3) закрытые данные. Этот класс данных предполагает, что к рукописям никто как из внешних пользователей сайта (учёных), так и из внутренних пользователей (администраторов базы данных) не будет иметь доступа. Обычно под этим классом данных хранят пароли пользователей на сайте, что минимизирует риск их взлома или передачи третьим лицам недобросовестными администраторами. На наш взгляд, несанкционированное проникновение в дневники и частные письма учёного есть такой же взлом, равносильный взлому пароля пользователя на сайте, поскольку у злоумышленника (т. н. «хакера») появляется доступ к мыслям и замыслам учёного. Необходимость сокрытия персональных данных, в том числе, устанавливается в ст. 138 УК РФ «Нарушение тайны переписки, телефонных переговоров, почтовых, телеграфных или иных сообщений» [24].

Таблица 4. Таблица классификации данных, с точки зрения обеспечения их конфиденциальности

№	Класс данных	Получатель доступа	Пример общего класса	Пример реализации класса в программе
1	Открытые	Учёные, администраторы базы данных	Имя и фамилия пользователя	Опубликованные рукописи
2	Полузакрытые	Администраторы базы данных	Email пользователя	Черновики рукописей
3	Закрытые	–	Пароль на сайте	Личные дневники и частная переписка

Помимо вышеперечисленного, важно продумать механизм, при котором потенциальные биографы смогут получать право на ознакомление с рукописями, дневниками и частными письмами учёных. На наш взгляд, будут уместны следующие основания для предоставления права на ознакомление с ними:

1) право беспрекословного доступа к рукописям учёного должны иметь члены его семьи при условии, что учёный не указал лицо, обращающееся за получением рукописей, в списке исключений в разделе наследников рукописей по тем или иным причинам. Например, учёный мог заподозрить того или иного члена семьи в возможном халатном обращении с рукописями после ухода учёного из жизни;

2) потенциальный биограф должен отправить администрации программы документ, удостоверяющий его потенциальные возможности выступить в роли биографа. Данным документом могут быть:

– **диплом о высшем образовании по гуманитарной специальности**. Если обратить внимание на то, какой именно тип образования имеют авторы монографий, посвящённых биографиям выдающихся индивидов, то это, как правило, представители гуманитарных специальностей (педагоги, философы, социологи и др.), в рамках которых человек занимает центральное место;

– либо должны быть отправлены примеры самостоятельно написанных или изданных художественных книг, что, с высокой долей вероятности, предполагает наличие у запрашивающего закреплённых навыков письменного изложения мыслей. Хотя, на наш взгляд, научное образование и, как следствие, поставленные навыки работы с научными источниками, а также знание основ источниковедения дают значительно больший приоритет для передачи электронных копий рукописей для изучения, поскольку авторы художественных книг могут принести пользу для науки в части её популяризации.

Если лицо, обращающееся за рукописями, является представителем точных или естественных наук, а идеи учёного относились к гуманитарным направлениям, то тогда доступ ему может быть предоставлен только при кооперации с представителем гуманитарных наук.

Подобного рода проверочные действия считаем важными для того, чтобы минимизировать риск некачественно написанных биографических работ о жизни учёных. Стоит отметить, что данное решение должно быть в дальнейшем подтверждено статистическими данными. Например, можно взять массив со 100 биографическими монографиями, опубликованными и прошедшими рецензирование (например, из серии книг «Жизнь замечательных людей», или ЖЗЛ, издательства «Молодая гвардия», которые изучались автором при написании настоящей статьи), подобранными с соблюдением принципа рандомизация, выписав их названия в первом столбце условной таблицы, а во втором указав специальности тех, кто являлся их автором. Таким образом, будет статистически определено, представителям каких специальностей стоит назначать более высокий приоритет при предоставлении прав на написание биографии того или иного учёного.

НЕОБХОДИМОСТЬ В РАЗРЕШЕНИИ «ПРОБЛЕМЫ СТОЛЕТИЯ»

Проблема столетия – это понятие, самостоятельно введенное в рамках данной статьи, которое характеризует одну из наиболее сложных задач, с которой предстоит столкнуться на этапе сопровождения программы, после этапа её конструирования для обеспечения её непрерывного функционирования в будущем.

Дело в том, что текущие системы управления базами данных (СУБД) не гарантируют, что хранение тех или иных данных будет возможно на протяжении десятков, а тем более сотен лет. Поскольку создатели систем управления базами данных – это, как правило, коммерческие структуры, то, как большинство коммерческих структур, они подвержены флуктуациям экономического рынка, вследствие чего могут в любой момент перестать функционировать на нём, например, в случае банкротства, смены руководства компании и т. д., а исходная концепция программы «История гениального открытия» предполагает, что все работы живущих в настоящее время в России Ломоносовых, Циолковских и Бехтеревых спустя 100 лет будут оставаться доступными для изучения сторонними исследователями.

Помимо прочего, уязвимость программы, с точки зрения её зависимости от времени, заключается в том, что рано или поздно из жизни уйдут не только пользователи программы, но и её главный конструктор, что ставит вопрос о необходимости введения элемента преемственности в процесс управления жизненным циклом программы. В связи с этим наиболее очевидной на 2019 г. видится идея конструирования и управления программой в формате Open Source [11, с. 125], то есть посредством предоставления доступа к исходному коду программы другим разработчикам, проявляющим интерес к использованию достижений программной инженерии на благо российской науки и ускорения её развития благодаря им.

Проблема столетия отчасти может быть решена посредством использования одновременно множества серверов, купленных сразу у нескольких наиболее крупных систем управления базами данных, что позволит настроить синхронное резервное копирование рукописей на них. Данный подход является одним из примеров т. н. «диверсификации рисков», которую часто используют инвесторы при хранении денежных капиталов. Поскольку автор убежден, что результат мыслительной деятельности учёного – это тоже некоторый капитал, отличающийся лишь своей интеллектуальной, идеальной направленностью, то и забота о данном интеллектуальном капитале должна быть обеспечена повышенным уровнем защищённости в процессе управления рисками.

ЗАКЛЮЧЕНИЕ

- Подведём некоторые итоги. Программа «История гениального открытия»:
- ориентирована, в основном, на учёных и сохранение научного и культурного капиталов, оставляемых ими после смерти;
 - по мере накопления больших данных, составленных из массивов результатов деятельности множества учёных (публикаций, монографий, учебников и др.), позволит изучить закономерности становления научных теорий и количественно дополнить теорию Т. Куна о стадиях осуществления «научных революций»;
 - позволит пролонгировать сохранившиеся научные работы во времени посредством их посмертной передачи специализированным структурам, например, научно-исследовательским институтам (НИИ);
 - позволит интегрировать современные достижения в области программной инженерии (веб-приложения, работающие полностью в браузерах, криптографию и т. д.) с научной предметной областью;
 - стандартизирует процесс работы со вторичными источниками и их перекодировкой в машинописные данные, что упростит процесс их понимания, минимизируя вероятность искажений или ошибок в случае их перевода в машинописный вид вследствие особенностей письменного стиля каждого отдельно взятого учёного;
 - будет аккумулировать первичные данные (дневниковые записи, частные письма, отфильтрованную историю посещений в интернете и т.д.) для их дальнейшего изучения и использования будущими биографами.

Благодарности

Выражаем благодарность ректору Института непрерывного профессионального образования (г. Москва), проф. С.В. Чернову. Во многом именно благодаря двум его трактатам, посвящённым проблематике гениальности – «История гениальности в России. Книга первая: Университет всея Руси. Монография» и «Книга о гениальности. Т. I: Человеческий гений: Природа. Сущность. Становление. Монография» – появился замысел создания программного обеспечения, описанного в настоящей публикации, и под его научным руководством была

сформулирована концепция «Истории гениального открытия». Также благодарим педагогический коллектив ИТИС КФУ (г. Казань), Альма-матер, в стенах которой нами совершались первые шаги в области изучения больших данных. Именно большие данные должны составлять основу для выявления количественных тенденций построения научных теорий.

СПИСОК ЛИТЕРАТУРЫ

1. *Альтшуллер Г.С., Верткин И.М.* Как стать гением: Жизн. стратегия творч. личности. Мн.: Беларусь, 1994. 479 с.
2. *Вебер М.* Избранное: Протестантская этика и дух капитализма. 2-е изд., доп. и испр. М.: «Российская политическая энциклопедия» (РОССПЭН), 2006. 656 с. (Серия «Книга света»).
3. *Гиппократ: Избранные книги // Пер. с греч. проф. В.И. Руднева; Ред., вступ. статьи и прим. проф. В.П. Карпова.* М.: Гос. изд-во биол. и мед. лит-ры, 1936. 736 с.
4. Наследование // «Гражданский кодекс Российской Федерации (часть третья)» от 26.11.2001 N 146-ФЗ (ред. от 18.03.2019). Статья 1110.
5. Охраняемые результаты интеллектуальной деятельности и средства индивидуализации // «Гражданский кодекс Российской Федерации (часть четвертая)» от 18.12.2006 N 230-ФЗ (ред. от 18.07.2019). Ст. 1225, п. 1.
6. Срок действия исключительного права на произведение // «Гражданский кодекс Российской Федерации (часть четвертая)» от 18.12.2006 N 230-ФЗ (ред. от 18.07.2019). Статья 1281.
7. *Данилевский И.Н. и др.* Источниковедение: Теория. История. Метод. Источники российской истории: Учеб. пособие/И.Н. Данилевский, В.В. Кабанов, О.М. Медушевская, М.Ф. Румянцева. М.: Российск. гос. гуманит. ун-т, 1998. 702 с.
8. *Демин В.Н.* Циолковский. М.: Молодая гвардия, 2005. 323[13] с.: ил.
9. *Кириллина Л.В.* Бетховен. Жизнь и творчество: в 2 т. Т. II / Л.В. Кириллина. М.: Научно-издательский центр «Московская консерватория», 2009. 596 с.
10. *Кун Т.* Структура научных революций: Пер. с англ. / Т. Кун; Сост. В.Ю. Кузнецов. М.: ООО «Издательство АСТ», 2003. 605, [3] с.

11. *Липаев В.В.* Программная инженерия. Методологические основы. Учебник. М.: ТЕИС, 2006. 608 с.

12. *Михед П.В.* Гоголь и Мериме: «Заметки о Мериме» в контексте духовно-эстетических исканий 1840-х гг. // Имагология и компаративистика. 2015. №2 (4).

13. *Мостинская А.Ю., Бодрихин Н.Г.* Сергей Капица: Человек, который отвечал на любой вопрос / Алла Мостинская, Николай Бодрихин. М.: Молодая гвардия, 2015. 350[2] с.: ил.

14. *Петухова С.А.* Библиография жизни и творчества П.И. Чайковского. Указатель литературы, вышедшей на русском языке за 140 лет (1866–2006) / Авт.-сост. С.А. Петухова. М.: Государственный институт искусствознания, 2014. 856 с.

15. Программа «Voyant Tools». Результат контент-анализа монографии проф. С.В. Савельева «Гениальность и изменчивость». URL: <https://voyant-tools.org/?corpus=0fb4d4746c68056dcba0c3688512b521> (дата обращения: 15.10.2019).

16. Программа «Voyant Tools». Результат контент-анализа монографии проф. С.В. Чернова «Книга о гениальности. Т. 1: Человеческий гений: Природа. Сущность. Становление. Монография». URL: <https://voyant-tools.org/?corpus=8d2e80989a9319cda04a8d7b885a74c3> (дата обращения: 15.10.2019).

17. Рабочая книга социолога / Под общ. ред. и с предисл. Г.В. Осипова. Изд. 5-е. М.: Книжный дом «ЛИБРОКОМ», 2009. 480 с.

18. Российская академия наук. Устав РАН: предмет, цели и виды деятельности, основные задачи и функции Академии. URL: <http://www.ras.ru/about/rascharter/tasks.aspx> (дата обращения: 17.04.2019).

19. Российский союз писателей. URL: <https://rossp.ru> (дата обращения: 25.05.2020).

20. *Суриков И.Е.* Сократ / Игорь Суриков. М.: Молодая гвардия, 2011. 365[3] с.: ил.

21. *Суходольский Г.В.* Основы математической статистики для психологов: Учебник. СПб.: Издательство С.-Петербургского университета, 1998. 464 с.

22. Таненбаум Э., Остин Т. Архитектура компьютера. 6-е изд. СПб.: Питер, 2013. 816 с.: ил.

23. Тинькофф Журнал (Т—Ж). Как диверсифицировать валютную часть портфеля? URL: <https://journal.tinkoff.ru/ask/deposits-and-etfs/> (дата обращения: 24.01.2020).

24. Нарушение тайны переписки, телефонных переговоров, почтовых, телеграфных или иных сообщений // «Уголовный кодекс Российской Федерации» от 13.06.1996 N 63-ФЗ (ред. от 07.04.2020) (с изм. и доп., вступ. в силу с 12.04.2020). Статья 138.

25. Федеральный закон «О персональных данных» от 27.07.2006 N 152-ФЗ (последняя редакция).

26. Черниговская Т.В. Чеширская улыбка кота Шрёдингера: язык и сознание. М.: Языки славянской культуры, 2013. 448 с.

27. Чернов С.В. История гениальности в России. Книга первая: Университет всея Руси. Монография / С.В. Чернов. М.: Издательство Института Непрерывного Профессионального Образования, 2017. 422 с.

28. Чернов С.В. Книга о гениальности. Т. 1: Человеческий гений: Природа. Сущность. Становление. Монография / С.В. Чернов, Воронеж. М.: АНО «Институт духовной культуры и свободного творчества», 2010. 562 с.

29. Шаталов М.А., Мычка С.Ю. Диверсификация как метод снижения финансовых рисков в условиях нестабильности экономики // Территория науки. 2014. №6.

30. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке С. М.: Издательство ТРИУМФ, 2003. 816 с.: ил.

31. Galarnyk M. The normal Gaussian distribution. URL: <https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2> (дата обращения: 01.02.2020).

32. GitHub. Global company that provides hosting for software development version control using Git. URL: <https://github.com/about> (дата обращения: 01.02.2020).

33. GitHub. The History of Genius Discovery software. URL: <https://github.com/R-Mosolov/history-gd> (дата обращения: 12.12.2020).

34. Wired. Lord of the Files: How GitHub Tamed Free Software (And More). URL: <https://www.wired.com/2012/02/github-2/> (дата обращения: 16.06.2020).

THE HISTORY OF GENIUS DISCOVERY SOFTWARE

R. V. Mosolov

Kazan (Volga region) Federal University, Kazan;

Schematus (USA), 54/1, Pushkina St., Kazan, 420015, Russian Federation

Abstract

This article description the conception of History of Genius Discovery (History GD) software. The software has few similarities with GitHub software that have got wide famous at the professional developer's community. The software appealed to solve two main science issues. History GD will save science and cultural heritage of Russian scientists and accumulate initial data for measuring tendencies of science theorems formation. The last will give the probability for appending The Structure of Scientific Revolutions by Thomas S. Kuhn by using numeric big data. Also, the software will minimise probability of losing scientific manuscripts by reason of scientists deaths. Software engineering, sociology, philosophy, law, and history are five scientific directions that are used as base for creating this software. The idea of creation have got at Kazan Federal University when we learned Big Data Science.

Keywords: *The History of Genius Discovery, History GD, scientific heritage, cultural heritage, genius patterns, scientific software, software for scientists, GitHub for scientists.*

REFERENCES

1. *Al'tshuller G.S., Vertkin I.M. Kak stat' geniem: Zhizn. strategija tvorch. lichnosti. Mn.: Belarus', 1994. 479 s.*
2. *Veber M. Izbrannoe: Protestantskaja jetika i duh kapitalizma. 2-e izd., dop. i ispr. M.: «Rossijskaja politicheskaja jenciklopedija» (ROSSPJeN), 2006. 656 s. (Serija «Kniga sveta»).*

3. *Gippokrat: Izbrannye knigi* // Per. s grech. prof. V.I. Rudneva; Red., vstup. stat'i i prim. prof. V.P. Karpova. M.: Gos. izd-vo biol. i med. lit-ry, 1936. 736 s.

4. Nasledovanie // «Grazhdanskij kodeks Rossijskoj Federacii (chast' tret'ja)» ot 26.11.2001 N 146-FZ (red. ot 18.03.2019). Stat'ja 1110.

5. Ohranjaemye rezul'taty intellektual'noj dejatel'nosti i sredstva individualizacii // «Grazhdanskij kodeks Rossijskoj Federacii (chast' chetvertaja)» ot 18.12.2006 N 230-FZ (red. ot 18.07.2019). St. 1225, p. 1.

6. Srok dejstvija iskljuchitel'nogo prava na proizvedenie // «Grazhdanskij kodeks Rossijskoj Federacii (chast' chetvertaja)» ot 18.12.2006 N 230-FZ (red. ot 18.07.2019). Stat'ja 1281.

7. *Danilevskij I.N. i dr. Istochnikovedenie: Teorija. Istorija. Metod. Istochniki rossijskoj istorii: Ucheb. posobie*/I.N. Danilevskij, V.V. Kabanov, O.M. Medushevskaja, M.F. Rumjanceva. M.: Rossijsk. gos. gumanit. un-t, 1998. 702 s.

8. *Demin V.N. Ciolkovskij*. M.: Molodaja gvardija, 2005. 323[13] s.: il.

9. *Kirillina L.V. Bethoven. Zhizn' i tvorcestvo: v 2 t. T. II* / L.V. Kirillina. M.: Nauchno-izdatel'skij centr «Moskovskaja konservatorija», 2009. 596 s.

10. *Kun T. Struktura nauchnyh revoljucij: Per. s angl.* / T. Kun; Sost. V.Ju. Kuznecov. M.: OOO «Izdatel'stvo ACT», 2003. 605, [3] s.

11. *Lipaev V.V. Programmaja inzhenerija. Metodologicheskie osnovy*. Uchebnik. M.: TEIS, 2006. 608s.

12. *Mihed P.V. Gogol' i Merime: «<3ametki o Merime>» v kontekste duhovno-jesteticheskikh iskanij 1840-h gg.* // *Imagologija i komparativistika*. 2015. #2 (4).

13. *Mostinskaja A.Ju., Bodrihin N.G. Sergej Kapica: Chelovek, kotoryj otvechal na ljuboj vopros* / Alla Mostinskaja, Nikolaj Bodrihin. M.: Molodaja gvardija, 2015. 350[2] s.: il.

14. *Petuhova S.A. Bibliografija zhizni i tvorcestva P.I. Chajkovskogo. Ukazatel' literatury, vyshedshej na ruskom jazyke za 140 let (1866–2006)* / Avt.-sost. S.A. Petuhova. M.: Gosudarstvennyj institut iskusstvoznaniya, 2014. 856 s.

15. Programma «Voyant Tools». Rezul'tat kontent-analiza monografii prof. S.V. Savel'eva «Genial'nost' i izmenchivost'». URL: <https://voyant-tools.org/?corpus=0fb4d4746c68056dcba0c3688512b521>

16. Programma «Voyant Tools». Rezul'tat kontent-analiza monografii prof. S.V. Chernova «Kniga o genial'nosti. T. 1: Chelovecheskij genij: Priroda. Sushhnost'. Stanovlenie. Monografija».

URL: <https://voyant-tools.org/?corpus=8d2e80989a9319cda04a8d7b885a74c3>

17. Rabochaja kniga sociologa / Pod obshh. red. i s predisl. G. V. Osipova. Izd. 5-e. M.: Knizhnyj dom «LIBROKOM», 2009. 480 s.

18. Rossijskaja akademija nauk. Ustav RAN: predmet, celi i vidy dejatel'nosti, osnovnye zadachi i funkcii Akademii.

URL: <http://www.ras.ru/about/rascharter/tasks.aspx>

19. Rossijskij sojuz pisatelej. URL: <https://rossp.ru>

20. *Surikov I.E. Sokrat / Igor' Surikov*. M.: Molodaja gvardija, 2011. 365[3] s.: il.

21. *Suhodol'skij G.V. Osnovy matematicheskoj statistiki dlja psihologov: Uchebnik*. SPb.: Izdatel'stvo S.-Peterburgskogo universiteta, 1998. 464 s.

22. *Tanenbaum Je., Ostin T. Arhitektura komp'jutera*. 6-e izd. SPb.: Piter, 2013. 816 s.: il.

23. Tin'koff Zhurnal (T—Zh). Kak diversificirovat' valjutnuju chast' portfelja? URL: <https://journal.tinkoff.ru/ask/deposits-and-etfs/> (data obrashhenija: 24.01.2020).

24. Narushenie tajny perepiski, telefonnyh peregovorov, pochtovyh, telegrafnyh ili inyh soobshhenij // «Ugolovnyj kodeks Rossijskoj Federacii» ot 13.06.1996 N 63-FZ (red. ot 07.04.2020) (s izm. i dop., vstup. v silu s 12.04.2020). Stat'ja 138.

25. Federal'nyj zakon «O personal'nyh dannyh» ot 27.07.2006 N 152-FZ (poslednjaja redakcija).

26. *Chernigovskaja T.V. Cheshirskaja ulybka kota Shrijodingera: jazyk i soznanie*. M.: Jazyki slavjanskoj kul'tury, 2013. 448 s.

27. *Chernov S.V. Istorija genial'nosti v Rossii. Kniga pervaja: Universitet vseja Rusi. Monografija / S. V. Chernov*. M.: Izdatel'stvo Instituta Nepreryvnogo Professional'nogo Obrazovanija, 2017. 422 s.

28. *Chernov S.V. Kniga o genial'nosti. T. 1: Chelovecheskij genij: Priroda. Sushhnost'. Stanovlenie. Monografija / S.V. Chernov, Voronezh*. M.: ANO «Institut duhovnoj kul'tury i svobodnogo tvorchestva», 2010. 562 c.

29. *Shatalov M.A., Mychka S.Ju.* Diversifikacija kak metod snizhenija finansovyh riskov v uslovijah nestabil'nosti jekonomiki // Territorija nauki. 2014. #6.

30. *Shnajer B.* Prikladnaja kriptografija. Protokoly, algoritmy, ishodnye teksty na jazyke C. M.: Izdatel'stvo TRIUMF, 2003. 816 s.: il.

31. *Galarnyk M.* The normal Gaussian distribution. URL: <https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2>

32. GitHub. Global company that provides hosting for software development version control using Git. URL: <https://github.com/about>

33. GitHub. The History of Genius Discovery software. URL: <https://github.com/R-Mosolov/history-gd> (data obrashhenija: 12.12.2020).

34. Wired. Lord of the Files: How GitHub Tamed Free Software (And More). URL: <https://www.wired.com/2012/02/github-2/> (data obrashhenija: 16.06.2020).

СВЕДЕНИЯ ОБ АВТОРЕ



Роман Валерьевич МОСОЛОВ – бакалавр социологии, магистрант Института информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета, программный инженер в Schematus (США) (коворкинг «Телеграф Space»).

MOSOLOV Roman Valer'evich – Bachelor in Sociology, Master Student in Computer Science at Institute ITIS of Kazan Federal University, Software Engineer at Schematus (USA), Telegraph Space Coworking.

email: R.V.Mosolov@ya.ru

Материал поступил в редакцию 10 октября 2020 года

УДК 004.550

V МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ДЛЯ НАУК О ЗЕМЛЕ И ПРИЛОЖЕНИЯ ДЛЯ ГЕОЛОГИИ, ГОРНОЙ ПРОМЫШЛЕННОСТИ И ЭКОНОМИКИ. ITES&MP-2019»

В. В. Наумова¹

Федеральное государственное бюджетное учреждение науки «Государственный геологический музей им. В.И. Вернадского РАН»

¹naumova_new@mail.ru

Аннотация

Охарактеризованы материалы, представленные на V международной конференции «Информационные технологии для наук о Земле и приложения для геологии, горной промышленности и экономики. ITES&MP-2019». Названная конференция описывает результаты последних лет в следующих областях:

- открытый доступ к научным данным в области наук о Земле; особенности данных в науках о Земле: новые концепции и методы, инструменты их сбора, интеграции и обработки в различных информационных системах, в том числе в системах с интенсивным использованием данных;
- анализ данных и математическое моделирование природных процессов в науках о Земле: новые подходы. Эволюция классических ГИС-приложений;
- применение информационных технологий в области металлогении критических полезных ископаемых;
- социальные аспекты горно-геологической отрасли;
- прогнозные построения в области геологической разведки и землепользования;
- интеллектуальный анализ данных, извлечение фактов и знаний из научных публикаций. Тезаурусы, онтологии, концептуальное моделирование. Семантический веб, связанные данные. Сервисы. Семантическое структурирование контента. Применение в науках о Земле;
- применение методов и технологий дистанционного зондирования в

науках о Земле и горной промышленности: от спутников до беспилотных летательных аппаратов;

- информационные технологии для создания систем демонстрации и популяризации достижений в науках о Земле;
- приложения: прогноз месторождений, экологические риски, опасные природные явления, управление водными ресурсами, геотермальная энергия и др.

Ключевые слова: *информационные технологии, науки о Земле.*

В Государственном геологическом музее им. В.И. Вернадского РАН (рис. 1) 14–18 октября 2019 года прошла V Международная конференция «Информационные технологии для наук о Земле и приложения для геологии, горной промышленности и экономики. ITES&MP-2019».



Рисунок 1. Государственный геологический музей им. В.И. Вернадского РАН, г. Москва.

Конференция ITES&MP-2019 была образована в результате преобразования конференций ITES «Современные информационные технологии в науках о

Земле» (Магадан, 2008; Владивосток, 2010; Петропавловск-Камчатский, 2014; Южно-Сахалинск, 2016). Основным организатором этих конференций – Дальневосточный геологический институт Дальневосточного отделения (ДВО) РАН (Лаборатория информационных технологий) при партнерском участии других институтов РАН: Северо-Восточного комплексного научно-исследовательского института (СВКНИИ) им. Н.А. Шило ДВО РАН, Института вулканологии и сейсмологии ДВО РАН, Института автоматизации и процессов управления ДВО РАН, Института геологии и природопользования ДВО РАН, Геофизического центра РАН, Государственного геологического музея им. В.И. Вернадского РАН) и Mineral Prospectivity Conference (Orléans, France, 2017, организатор этой конференции – Геологическая служба Франции (BRGM) при партнерском участии других организаций) с намерением создать форум, отражающий современные проблемы организации данных, их исследования и анализа в науках о Земле. Преобразование конференций, названных выше, было организовано так, что были сохранены связи с ITES и Mineral Prospectivity Conference, а также сообщество участников ITES и Mineral Prospectivity Conference, сформированное в течение ряда лет успешной работы. Экспоненциальный рост данных, методов и систем обработки практически во всех областях наук о Земле, а также понимание роли, которую играют и будут играть информационные технологии в этих областях, были основной мотивацией для организации Конференции. Задачи проведения Конференции покрывают весь спектр вопросов, связанных с поиском и разведкой месторождений полезных ископаемых, рациональным природопользованием; сбором, обработкой, анализом данных; управлением и популяризацией знаний.

Организаторы Конференции: Российская академия наук, Академия горных наук РФ, Государственный геологический музей им. В.И. Вернадского РАН, Геологическая служба Франции.

Сопредседателями международного Программного комитета Конференции были: А.О. Глико, академик РАН, академик-секретарь Отделения наук о Земле РАН, г. Москва (Россия); Ю.Н. Малышев, академик РАН, Президент, Государственный геологический музей (ГГМ) им. В.И. Вернадского РАН, г. Москва (Россия); А.Ф. Морозов, заместитель руководителя, Федеральное Агентство по недропользованию РФ – Роснедра, г. Москва (Россия); Jean-Claude Guillaneau,

директор по георесурсам, Геологическая служба Франции (BRGM), Orléans (France). Заместителем председателя была В.В. Наумова (ГГМ РАН), ученым секретарем – И.С. Голубенко (СВКНИИ ДВО РАН). В состав Программного комитета конференции входили ведущие ученые, работающие в данном направлении: И.В. Бычков, академик РАН, Институт динамики систем и теории управления Сибирского отделения (СО) РАН им. В.М. Матросова / Иркутский государственный университет, г. Иркутск (Россия); А.Д. Гвишиани, академик РАН, Геофизический центр РАН, Национальный геофизический комитет, Комитет по системному анализу РАН, г. Москва (Россия); Е.И. Гордеев, академик РАН, Институт вулканологии и сейсмологии ДВО РАН, г. Петропавловск-Камчатский (Россия); Н.А. Горячев, член-корреспондент РАН, Северо-Восточный комплексный научно-исследовательский институт ДВО РАН, г. Магадан (Россия); А.М. Елизаров, д. ф.-м. н., профессор, Казанский федеральный университет, г. Казань (Россия); О.Л. Жижимов, д. т. н., Институт вычислительных технологий СО РАН, г. Новосибирск (Россия); В.А. Левин, академик РАН, Институт автоматизации и процессов управления ДВО РАН, г. Владивосток / Московский государственный университет, г. Москва (Россия); Л.В. Массель, д. т. н., Институт систем энергетики СО РАН, г. Иркутск (Россия); В.П. Потапов, д. т. н., Институт вычислительных технологий СО РАН, Кемеровский филиал, г. Кемерово (Россия); Г.М. Ружников, д. т. н., Институт динамики систем и теории управления СО РАН им. В.М. Матросова, г. Иркутск (Россия); В.А. Серебряков, д. ф.-м. н., профессор, Вычислительный центр ФИЦ «Информатика и управление» РАН, г. Москва (Россия); А.М. Федотов, член-корреспондент РАН, Институт вычислительных технологий СО РАН, г. Новосибирск (Россия); Л.Е. Чесалов, д. т. н., ФГБУ «Гидроспецгеология», г. Москва (Россия); Е.Н. Черемисина, д. т. н., «ВНИИгеосистем», г. Москва/ Университет «Дубна», г. Дубна (Россия); Dr. Cassard Daniel, BRGM, Orléans (France); Mr. Jean-Marc Trouillard, BRGM, Orléans (France); Prof. John Carranza, University of KwaZulu-Natal, Durban (South Africa); Prof. Klaudia Oleschko, Universidad Nacional, Autonoma de México, Queretaro (México); Dr. Marko Komac, Assoc. Prof., D. Sc. Geol., European Federation of Geologist Board Member – External Relations Officer / Geological Survey of Slovenia, Ljubljana (Slovenia); Dr. Michel Bonnemaïson, E-Mines, Orléans (France); Dr. Phung Van Phach, Institute for Marine Geology & Geophysics, Vietnam-

ese Academy of Science & Technology (Vietnam); Dr. Reimar Seltmann, Natural History Museum, London (Great Britain); Dr. Warren J. Nokleberg, U.S. Geological Survey, Menlo Park, California (USA); Dr. Yuei-An Liou, National Central University, Taiwan Group on Earth Observations, (Taiwan); Dr. Andreas Barth, Beak Consultants GmbH (Germany); Dr. Charles Gumiaux, Institut des Sciences de la Terre (ISTO), Orleans University (France); Prof. Danielle Ziebelin, Laboratoire d'Informatique de Grenoble (LIG), Grenoble-Alpes University (France).

Конференция прошла при поддержке Посольства Франции в России и спонсорской поддержке компаний Микромайн, Intel, Softline и Международных организаций IAGOD, SGA и SEG. Партнером Конференции была Группа компаний СканЭкс.

Основные направления работы Конференции:

- Открытый доступ к научным данным в области наук о Земле.
 - Особенности данных в науках о Земле: новые концепции и методы, инструменты их сбора, интеграции и обработки в различных информационных системах, в том числе в системах с интенсивным использованием данных.
 - Анализ данных и математическое моделирование природных процессов в науках о Земле: новые подходы. Эволюция классических ГИС-приложений.
 - Применение информационных технологий в области металлогении критических полезных ископаемых; социальные аспекты горно-геологической отрасли; прогнозные построения в области геологической разведки и землепользования.
 - Интеллектуальный анализ данных, извлечение фактов и знаний из научных публикаций. Тезаурусы, онтологии, концептуальное моделирование. Семантический веб, связанные данные. Сервисы. Семантическое структурирование контента. Применение в науках о Земле.
 - Применение методов и технологий дистанционного зондирования в науках о Земле и горной промышленности: от спутников до беспилотных летательных аппаратов.
 - Информационные технологии для создания систем демонстрации и популяризации достижений в науках о Земле.
-

- Приложения: прогноз месторождений, экологические риски, опасные природные явления, управление водными ресурсами, геотермальная энергия и др.

В Конференции приняли участие более 100 ученых из 8 стран: России, Франции, Армении, Казахстана, Южно-Африканской республики, Египта, Танзании, Шри-Ланки, в том числе из 15 городов РФ: Москвы, Санкт-Петербурга, Новосибирска, Красноярска, Иркутска, Владивостока, Магадана, Ханты-Мансийска, Перми, Апатит, Барнаула, Кемерово, Фрязино, Архангельска, Твери.

За три дня Конференции было заслушано и обсуждено 70 докладов по различным темам, связанным с ее тематикой.

Открыл Конференцию Ю.Н. Малышев, сопредседатель Международного программного комитета Конференции, президент Государственного Геологического музея им. В.И. Вернадского РАН, д. т. н., академик РАН, г. Москва (Россия).

На Пленарной сессии (рис. 2) были озвучены и обсуждены следующие доклады (<http://video.sgm.ru/records/20/1>):



Рисунок 2. Сопредседатели Пленарной сессии С.В. Черкасов и В.В. Наумова, Государственный геологический музей им. В. И. Вернадского РАН, Москва.

А.Д. Гвишиани, академик РАН, научный руководитель Геофизического центра РАН, г. Москва, представил доклад «Системный анализ в изучении Арктики».

В докладе Л.Е. Чесалова, ФГБУ Геоспецгеология, г. Москва, «Государ-

ственный репозиторий данных: вызовы и решения» были рассмотрены вопросы создания государственных информационных ресурсов. Освоение минерально-сырьевой базы мира в основном ведется крупными транснациональными корпорациями. Создавая свои информационные ресурсы, эти организации видят в них инструмент извлечения прибыли и конкурентной борьбы. Этим обусловлены строгие требованиями к конфиденциальности корпоративной информации. Такой подход неизбежно входит в конфликт с интересами стран, провинций и т. п. Целью государства является рациональное использование природных ресурсов. Государства и регионы контролируют не только финансовые результаты индустрии, но и собирают информацию об источнике этого дохода – природных ресурсах. Именно эти сведения будут основой как материального благополучия, так и безопасности среды обитания. Цифровая трансформация добывающей отрасли невозможна без надежных источников данных. При этом любая информация имеет значение и должна быть сохранена, поскольку имеющийся уровень знаний не может дать точного ответа о возможных потребностях завтрашнего дня. Рассматривается опыт разных стран по созданию и функционирования национальных банков информации. Подготовлены рекомендации для создания единого фонда информации о недрах.

Доклад «Система обеспечения работ по геологической поддержке недропользования (Собр Роснедра): состояние, перспективы развития» был озвучен Е.Н. Черемисиной, руководителем подразделения «Геоинформатика» «ВНИИ-Геосистем» ФГБУ «Всероссийский научно-исследовательский геологический нефтяной институт», г. Москва.

О разработке веб-ГИС среды для Армении, предназначенной для хранения, организации, визуализации, анализа и обмена геопространственными данными, которые используются в решении задач в науках о Земле, рассказал А.А. Аракелян из Института геологических наук Национальной академии наук Армении, г. Ереван (Республика Армения) в докладе «Открытый доступ к пространственным данным и знаниям в науках о Земле через виртуальную геонаучную среду для Армении». Среда Web-GIS разрабатывается сотрудниками лаборатории геоинформатики Института геологических наук (ИГН НАН) и названа «Виртуальная геонаучная среда для Армении» (VGSE, доступна по

адресу <http://vgse.geology.am>). Геопространственные данные различной тематики, накопленные в лаборатории в результате научных работ и проектов в течение более 20 лет, развитие веб-ГИС-систем и приложений и инициатив открытого доступа в последнее десятилетие, побудили авторов этой работы создать данную среду для научного сообщества в Армении. Авторы предполагают, что в результате обмена данными между различными организациями и исследователями, работающими в сфере наук о Земле, содержание и темы VGSE будут расширяться, и это может послужить основой для новых возможностей сотрудничества.

В докладе В.В. Наумовой с соавторами «Основные подходы к разработке территориально-распределенной информационно-аналитической геологической среды» из Государственного геологического музея им. В.И. Вернадского, г. Москва, была представлена разработка и адаптация методов и технологий для обработки и анализа территориально распределенной разнотипной геологической информации и сервисов ее обработки. На основе созданных подходов, разработанных методов и технологий, проведенного проектирования реализована основа информационно-аналитической среды для поддержки и сопровождения научных исследований в геологии, осуществляющая доступ к разнородной территориально распределенной геологической информации с использованием специализированных сервисов ее обработки. Авторами предполагается, что разработанная платформа управления тематическими сервисами обработки и анализа, входящая в состав информационно-аналитической среды, обеспечит пользователям доступ к хранилищам современных наукоемких алгоритмов и вычислительным ресурсам, необходимым для оперативной обработки больших массивов разнотипных геологических данных.

С. Дельчини, BRGM, Французская геологическая служба (г. Орлеан, Франция) в своем докладе с соавторами рассмотрел современные возможности сканера кернов SOLSA: автоматизированного мультифункционального аналитического инструмента для анализа буровых кернов пород со слоистой структурой.

Комбинированный анализ в режиме реального времени на буровых колонках позволяет эффективно, рентабельно и устойчиво улучшить разведку и добычу рудных месторождений, что приводит к оптимальному определению

геометаллургических параметров и более уверенным геомоделям и экономической оценке. В этом контексте проект SOLSA EU-H2020 (www.solsa-mining.eu) направлен на разработку экспертной системы, объединяющей ультразвуковое бурение с оперативным аналитическим сканером керна в реальном времени для работы в никелевых латеритовых средах и других типах руды. Система сканирования керна впервые объединяет систематический морфологический, текстурный, минералогический и химический анализ кернов с использованием профилометра, RGB и инфракрасных гиперспектральных камер (VNIR, SWIR) и рентгенофлуоресцентного спектрометра (XRF).

В докладе Л.В. Массель с соавторами из Института систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск, «Методы и информационные технологии оценки влияния энергетики на геоэкологию с учетом качества жизни» были рассмотрены методы и технологии оценки влияния энергетики на геоэкологию с учетом качества жизни, разрабатываемые в рамках поддерживаемого фондом ЕАПИ (Евро-Азиатский фонда поддержки научных исследований) международного проекта, выполняемого совместно с учеными из Беларуси и Армении. Были описаны результаты онтологического инжиниринга предметной области и методика перехода от онтологических моделей к моделям данных и знаний. Предложено при оценке влияния энергетики на геоэкологию региона учитывать качество жизни населения, с включением в интегральный показатель качества жизни показателей, характеризующих обеспеченность населения энергоресурсами требуемого качества и показателей, характеризующих загрязнение окружающей среды. Для этой цели предложено использовать когнитивное моделирование, приведены примеры когнитивных карт. Рассмотрена архитектура веб-ориентированной информационно-аналитической системы (WIS), интегрирующей уровни математического и семантического моделирования, уровня баз данных и баз знаний и геоинформационных систем. Предложена технология интеллектуальной поддержки принятия решений в области энергетики и экологии, направленных на повышение интегрального показателя качества жизни.

Эммануэль Джон М. Карранса, Университет Квазулу-Натал (Южная Африка) с соавторами из Института наук о Земле, Государственный университет Кампинас (Бразилия) представил результаты работ по анализу пространственных

шаблонов минеральных месторождений: примеры из минеральной провинции Каражас, Бразильский Амазон (рис. 3).



Рисунок 3. Эммануэль Джон Карранса, Университет Квазалу-Натал, Дурбан, ЮАР.

Формирование и пространственная структура месторождений полезных ископаемых обусловлены взаимодействием различных процессов, регулируемых геологической средой, в которой они существуют. Исследования в этом аспекте постепенно продемонстрировали, что некоторые из этих процессов имеют фрактальное распределение (то есть они являются масштабно-инвариантными). Это означает, что, несмотря на ощутимую сложность пространственной структуры месторождений полезных ископаемых в различных географических масштабах, существует регулярная простран-

ственная структура. При правильном анализе этот регулярный пространственный паттерн может быть полезным для различных смежных областей. Охарактеризована пространственная структура залежей оксида железа-меди-золота (IOCG) в известной минеральной провинции Каражас. В частности, пространственная структура месторождений IOCG исследуется в региональном масштабе (месторождения полезных ископаемых на региональных картах), локальном масштабе (рудные тела на картах рудников) и в микромасштабном масштабе (минералы руды в тонких разрезах). Продемонстрировано, что пространственная структура рудных минералов в микроуровне в основном неслучайная, показывает фрактальное распределение и четкие тренды, которые имитируют таковые для рудных тел на локальном уровне и месторождений полезных ископаемых на региональном уровне. Таким образом, значительная часть фрактальности пространственной структуры месторождений полезных ископаемых обусловлена структурным контролем минерализации, поскольку структуры являются преобладающим контролем потока жидкости. Отсюда следует, что основным фактором, контролирующим наблюдаемую фрактальную природу пространственной структуры месторождений полезных ископаемых, является проницаемость, которая по существу связана с многомасштабными структурами. Эти результаты дают представление о фрактальной природе контроля месторождений полезных ископаемых, а также новых многомасштабных стратегиях для изучения структурного контроля минерализации и инновационных подходов к картированию минерального потенциала.

О системе оперативной обработки спутниковых данных дистанционного зондирования в Красноярском научном центре СО РАН доложил О.Е. Якубайлик из Института вычислительного моделирования Сибирского отделения Российской академии наук, г. Красноярск (рис. 4). В докладе рассмотрены программно-технологическое обеспечение, информационно-вычислительная инфраструктура регионального центра ДЗЗ в Красноярске. Реализация аппаратно-программного обеспечения строится по модульному принципу как совокупность функционально самостоятельных компонентов. Логическими единицами системы являются группы серверов обработки спутниковых данных: 1) приема, 2) предварительной

обработки, 3) тематической обработки (стандартной и для конечных пользователей). Программное обеспечение реализуется в сервис-ориентированной архитектуре, в основном – на основе открытого и свободного программного обеспечения. Интерфейс пользователя создается на основе стандартного веб-браузера, с использованием технологий геоинформационных веб-систем и геопорталов, концепции инфраструктуры пространственных данных.



Рисунок 4. О.Э. Якубайлик, Институт вычислительного моделирования СО РАН, ФИЦ КНЦ СО РАН, г. Красноярск.

Основное внимание на текущем этапе уделяется созданию интерактивных веб-сервисов оперативного представления и экспресс-анализа спутниковой информации в «квазиреальном» режиме времени (доступность в течение часа после приема). Для спутниковых данных основных низкоорбитальных метеорологических спутников (Terra/Aqua/Suomi NPP/NOAA-20) автоматически формируется коллекция мультимасштабных композитных изображений с популярными комбинациями спектральных каналов, значимыми характеристиками (вегетаци-

онные индексы, яркостная температура и проч.). Также формируются прикладные специализированные сервисы углубленной обработки спутниковых данных.

Доклад А.А. Любушина из Института физики Земли им. О.Ю. Шмидта РАН, г. Москва, был посвящен окружающему шуму Земли и трендам глобальной сейсмической опасности.

Окружающий шум Земли означает постоянные вариации параметров сейсмического шума и тремора земной поверхности, измеренных с помощью космической геодезии. Источником этих случайных колебаний являются воздействия на земную кору со стороны атмосферы и океана, а также внутренние процессы в земных оболочках, в том числе предвидение крупных геологических катастроф. В настоящее время число активных станций GPS, которые регистрируют смещения земной поверхности в трех направлениях с шагом по времени 5 минут, превышает 11 тысяч и постоянно растет. Количество широкополосных сейсмических станций в глобальной сети исчисляется сотнями. В докладе были представлены результаты анализа глобальных и региональных сейсмических шумов с использованием ряда нелинейных статистических данных, оцениваемых в движущихся временных окнах. Набор статистических данных, используемых для изучения свойств сейсмического шума и дрожания земной поверхности, включает мультифрактальные и энтропийные свойства сигналов. Выделены ритмы изменения когерентности и трендов параметров сейсмического шума с периодом около 2,5 лет, предположительно вызванные резким нарушением регулярности вращения Земли в 2003 году.

Спонсоры Конференции представили новейшие разработки своих компаний.

Ольга Адрианова, менеджер по развитию бизнеса Intel SW, Россия/СЕЕ, представила программные решения Intel для HPC и наук о Земле.

В докладе М.А. Сергеевой (SCANEX Group (Россия) были рассмотрены прикладные геосервисы реального времени.

В настоящее время важнейшим вектором в развитии российской компании «СКАНЭКС» является переход от «сырьевой» концепции дистанционного зондирования, когда снимок и созданные на его основе стандартные продукты рассматриваются как ведущее направление бизнеса ДЗЗ, к сервисному подходу,

ориентированному на требования определенного сектора экономики или даже конкретного потребителя.

Оперативность предоставления результатов и многовариантность предлагаемых продуктов наряду с использованием открытых источников информации позволяет решать широкий круг задач, а также значительно минимизировать расходы на использование исходных данных. Вполне естественно, что удобство пользования сервисами в большинстве своем является одним из основных критериев выбора, и здесь выигрывают легко интегрируемые системы, в которых реализованы возможности создания продукта, ориентированного на решение задач пользователя. Очень важным условием являются непрерывное совершенствование алгоритмов и методов обработки данных, интеграция информационных потоков, а также формирование новых направлений их использования.

В докладе LLC MICROMINE RUS (Россия) «Роль ГГИС в создании цифрового двойника месторождения», озвученном Е.С. Шульга, были рассмотрены вопросы применения ГГИС для создания цифрового двойника месторождения, который подразумевает трехмерную блочную геологическую модель рудного тела и вмещающих пород, а также сеть разведочных и вскрывающих выработок. Описан процесс импорта и обработки исходных полевых данных в специализированном программном обеспечении Micromine с целью дальнейшей их интерпретации. В ходе обработки цифровой информации выполняются статистический и геостатистический анализы, выделение рудных интервалов с учетом кондиций, построение контуров рудных тел на продольных и поперечных разрезах с дальнейшим их преобразованием в трехмерную каркасную модель. Каркасная модель служит для создания блочной модели месторождения, на основании которой в дальнейшем выполняется подсчет запасов с категоризацией.

Сессия I. Особенности данных в науках о Земле: новые концепции и методы, инструменты их сбора, интеграции и обработки в различных информационных системах, в том числе в системах с интенсивным использованием данных. Открытый доступ к научным данным в области наук о Земле.

На этой сессии (рис. 5) были заслушаны и обсуждены доклады: И.М. Алёшина с соавторами (Институт физики Земли имени О.Ю. Шмидта РАН) «Центр агрегации геофизических данных "ИФЗ РАН"; В.Н. Ардасенова (LLC, SIE «FLATEK»)

«Основные пространственные данные (ОПД): аспект качественной оценки»; С.В. Булова с соавторами (Государственный геологический музей им. В.И. Вернадского РАН) «Геопортал «Металлогения» ГГМ РАН: структура и функциональность»; А. Одинцовой и А. Рыбкиной (Геофизический центр РАН) «Проект РОЗА ГИС: базы данных и веб-сервисы»; В.С. Еременко и В.В. Наумовой (Государственный геологический музей им. В.И. Вернадского РАН) «Вычислительно-аналитическая среда для обработки и анализа геологических данных»; В.К. Гусякова (Институт вычислительной математики и математической геофизики СО РАН) «Наблюдательные каталоги и базы данных в изучении природных катастроф: проблемы их составления, параметризации и представления»; В.В. Наумовой с соавторами (Государственный геологический музей им. В.И. Вернадского РАН) «Основные принципы разработки открытого доступа к фондовым данным государственного геологического музея им. В.И. Вернадского РАН»; К.А. Платонова (Государственный геологический музей им. В.И. Вернадского РАН) «Современные подходы в проектировании системы интеграции таблиц количественных данных в геологии»; А.А. Соловьёва с соавторами (Геофизический центр РАН) «Разработка веб-приложений для взаимодействия с данными наук о Земле»; Иванова С.Д. (Институт физики Земли имени О.Ю. Шмидта РАН, Москва (Россия) «Интерактивный реестр методов обработки данных для задач геохимии и петрологии»; Г.М. Ружникова с соавторами (Институт динамики систем и теории управления имени В.М. Матросова, СО РАН, г. Иркутск, Россия) «Сервис-ориентированная информационно-аналитическая система оценки влияния модели литосферы на динамические параметры колебаний скального грунта от землетрясений Южного Прибайкалья».

Сессия II. Анализ данных и математическое моделирование природных процессов в науках о Земле: новые подходы. Эволюция классических ГИС-приложений. Интеллектуальный анализ данных, извлечение фактов и знаний из научных публикаций. (<http://video.sgm.ru/records/20/2>, <http://video.sgm.ru/records/20/3>). Доклады этой секции показали широкий спектр применения различных математических методов, методов прикладной математики и геоинформационных технологий для решения задач в области наук о Земле: в области физики атмосферы; океанологии; в геологии; в экологии, при прогнозе чрезвычайных

природных ситуаций.



Рисунок 5. На одном из секционных заседаний. Слева направо: О.Ю. Кошель, Государственный геологический музей им. В.И. Вернадского РАН, г. Москва; Лафид Абделтиф, Французская геологическая служба, г. Орлеан, Франция; Сильван Дельчини, Французская геологическая служба, г. Орлеан, Франция; С.Д. Иванов, Институт физики Земли РАН, г. Москва; Е.П. Майсюк, Институт систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск.

Сессия III. Применение методов и технологий дистанционного зондирования в науках о Земле и горной промышленности: от спутников до беспилотных летательных аппаратов. (<http://video.sgm.ru/records/20/4>)



Рисунок 6. Д.М. Ермаков, Институт радиотехники и электроники РАН, г. Фрязино.

Все доклады этой секции можно разделить на две группы: разработка новых методов, технологий и систем обработки данных и непосредственно применение данных дистанционного зондирования для решения научных задач в области наук о Земле. На сессии были заслушаны доклады: «Беспилотные авиационные системы для комплексного бортового дистанционного зондирования как важнейший фактор трансформации геологоразведочной практики опыт 2014-2019 гг.» (А.В. Паршин); «Региональный удалённый мониторинг окружающей среды технологии и пользователи» (В.А. Левин, М.Г. Алексанина); «В поисках матрицы тектонических узлов севера российских плит по наземным и космическим данным» (З.Б. Чистова, Ю.Г. Кутинов); «Улучшение центра тропических циклонов с использованием вертикальных атмосферных профилей» (А.С. Еременко, К.С. Белоусова); «Развитие климатологической базы данных атмосфер-

ных рек» (Д.М. Ермаков, рис. 6.); «Использование метода главных компонент в анализе облачной структуры полярных циклонов по мультиспектральным спутниковым данным Modis Aqua/Terra» (Н.В. Федосеева, В.О. Лопуха); «Применение данных дистанционного зондирования (ДЗЗ) для картографирования структурных элементов, управляющих минерализацией золота на примере перспективных районов в Саха (Якутия) и Магаданской области (Россия)» (В.С. Муравьев, Н.В. Бондаренко, И.А. Гвоздева); «Решение проблемы быстрой оценки загрязнения воздуха, на основании геопортала ИКМ СО РАН» (А.А. Кадочников, А.В. Токарёв, О.Е. Якубайлик); «Визуализация в системе мониторинга космической погоды» (К.И. Холодков, С.Д. Иванов, И.М. Алёшин, В.А. Буров, Ю.П. Очелков).

Сессия IV. Применение информационных технологий в области металлогении критических полезных ископаемых; социальные аспекты горно-геологической отрасли; прогнозные построения в области геологической разведки и землепользования. (<http://video.sgm.ru/records/20/5>)

На этой сессии (рис. 7) были заслушаны и обсуждены доклады: Д.А. Клебанов «Цифровизация горной промышленности»; А. *Лахфид*, С. *Дельчини*, Т. *Баудин* «Использование геотермометрии методом рамановской спектроскопии углеродистого материала в горной промышленности»; Р. *Милло*, Э. *Глоаген*, Б. *Санджуан* «Изотопная геохимия и георесурсы лития: последние применения и будущие направления»; Р. *Милло*, Х. *Ибрагим*, Н. *Лафори*, Б. *Гусеро*, Дж. *Меллетон*, Т. *Конте*, Дж.П. *Гестем* «Гидрогеохимическое исследование золота в поверхностных водах с использованием интегративных и пассивных образцов: подтверждение концепции»; Е.П. *Масюк*, И.Ю. *Иванова*, А.К. *Ижбулдин* «Подход к оценке влияния функционирования энергетических объектов на геоэкологию прибрежных районов озера Байкал».

Организаторы Конференции полагают, что она будет способствовать повышению эффективности использования результатов научной деятельности органами исполнительной власти и организациями, связанными с контролем и управлением природными ресурсами, предупреждению и ликвидации последствий чрезвычайных природных и антропогенных ситуаций, проектно-исследовательскими работам. Организаторы надеются, что конференция послужит

стимулом для развития информационных технологий в научных исследованиях и образовании в области наук о Земле в России и будет способствовать привлечению молодежи в науку.



Рисунок 7. Сильван Дельчини и Мийо Ромен, Французская геологическая служба, г. Орлеан, Франция.

Молодым участникам Конференции, сделавшим лучшие доклады по версии ГК «СканЭкс», вручены лицензии на программное обеспечение компании.

Сайт Конференции: <http://ites2019.sgm.ru>. Заседания Конференции прошли в режиме прямой трансляции в интернет. Доклады участников Конференции опубликованы на видео-портале ГГМ РАН <http://video.sgm.ru/records/20/>.

Краткие тезисы и Программа Конференции опубликованы до начала Конференции [2, 3]. Избранные доклады Конференции опубликованы в CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073). Vol-2527 [1].

Программа Конференции включала в себя ряд экскурсий: по территории Московского Кремля и экспозициям Государственного геологического музея им. В.И. Вернадского РАН.

Конференция организована и проведена в рамках работ по Государственному заданию ФГБУН Государственного геологического музея им. В.И. Вернадского РАН по теме № 0140-2019-0005 «Разработка информационной среды интеграции данных естественнонаучных музеев и сервисов их обработки для наук о Земле».

СПИСОК ЛИТЕРАТУРЫ

1. CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073). Vol-2527. Edited by Vera V. Naumova, Aleksandr S. Eremenko.
2. Proceedings of the V International Conference, Moscow (Russia), 14–18 October 2019. Moscow: VNIIGeosystem, 2019.
3. Program of the V International Conference, Moscow (Russia), 14–18 October 2019. Moscow: VNIIGeosystem, 2019.

**V INTERNATIONAL CONFERENCE «INFORMATION TECHNOLOGIES
IN EARTH SCIENCES AND APPLICATIONS FOR GEOLOGY, MINING
AND ECONOMY. ITES&MP-2019»**

Vera V. NAUMOVA

Vernadsky State Geological Museum RAS, Moscow (Russia)

e-mail: Naumova_new@mail.ru

Abstract

The materials presented at the Conference describe the results of recent years in the following areas: Open access to scientific data and knowledge in Earth Sciences; Data peculiarities in Earth Sciences: new concepts and methods, tools for their collection, integration and processing in different information systems, including systems with intensive use of data; Data mining and mathematical simulation of natural processes in Earth Sciences. Evolution of classical GIS-applications in Earth Sciences; Application to Critical Raw Materials (CRM); social aspects of mining (e.g., the Social Licence to Operate [SLO]); predictive mapping and applications to exploration, landuse and search for extensions of known deposits; Intelligent data analysis, elicitation of facts and knowledge from scientific publications. Thesauruses, ontologies and conceptual modeling. Semantic WEB, linked data. Services. Content semantic structuring. Applications for geosciences, e.g., Ontology-based Dynamic Decision Graphs for Expert systems and decision-aid tools; Application of methods and technologies of the remote sensing in Earth Sciences: from satellites to unmanned aerial vehicles; Information technologies for demonstration and popularization of scientific achievements in Earth Sciences; Applications: environmental risks including mining wastes, natural hazards, water resource management, etc.

Keywords: Information technology, Earth sciences

REFERENCES

1. CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073). Vol-2527. Edited by Vera V. Naumova, Aleksandr S. Eremenko.
2. Proceedings of the V International Conference, Moscow (Russia), 14–18 October 2019. Moscow: VNIIGeosystem, 2019.

3. Program of the V International Conference, Moscow (Russia), 14–18 October 2019. Moscow: VNIIGeosystem, 2019.

СВЕДЕНИЯ ОБ АВТОРЕ



НАУМОВА Вера Викторовна – д. г.-м. н., г. н. с.,
зав. Научным отделом геолого-минералогического музея
им. В.И. Вернадского РАН, г. Москва, Россия

Vera V. NAUMOVA – prof., head of SGM scientific de-
partment, Vernadsky State Geological Museum RAS, Moscow
(Russia).

Email: Naumova_new@mail.ru

Материал поступил в редакцию 8 июля 2020 года

УДК 004.423.25

ОПИСАНИЕ КОНТЕКСТНО-СВОБОДНЫХ ГРАММАТИК В ФОРМАТЕ ДАННЫХ JSON ДЛЯ ГЕНЕРАТОРОВ СИНТАКСИЧЕСКИХ АНАЛИЗАТОРОВ

О. К. Осипов¹

Московский государственный технический университет им. Н. Э. Баумана

¹ osipovok@student.bmstu.ru

Аннотация

Рассмотрены варианты представления контекстно-свободных грамматик, предлагаемые средствами генерации синтаксических анализаторов. Приведён анализ существующих решений. Предложен новый формат описания грамматики. Дано описание грамматики в виде JSON-документа. Разработана концепция нового генератора, основанная на формате данных JSON для контекстно-свободных грамматик. Описана схема построения анализатора на основе концепции.

Ключевые слова: JSON-документ, контекстно-свободные грамматики, лексема, форма Бэкуса-Наура, дерево разбора, терминальные символы (токены), конечный детерминированный автомат, парсер, Parglare, ANTLR.

ВВЕДЕНИЕ

На сегодняшний день существует множество программных средств, которые строят синтаксический анализатор (именуемый также разборщиком или парсером) по данной им контекстно-свободной грамматике. Хотя большинство из них берет за основу дополненную нотацию Бэкуса-Наура [1] (сокр. на англ. ABNF, Augmented Backus Naur Form), некоторые добавляют свои расширения во входной формат грамматики, что ведёт к увеличению сложности и времени для его освоения. Кроме того, не все средства строго следуют каждому предписанию в стандарте ABNF.

В данной статье предложена концепция описания грамматик на основе формата данных JSON (JavaScript Object Notation – нотация объектов JavaScript) [2], [3]. Этот формат определяет структуру и возможное содержимое передава-

емых данных между приложениями в различных информационных системах (начиная от простого веб-сервиса и заканчивая системами управления баз данных и электронных библиотек). Его легко изучить и достаточно просто реализовать в рамках передачи данных между информационными системами. Он широко распространён во множестве современных веб-сервисов. Описано применение концепции к разработке генератора синтаксического разбора грамматик.

ОСОБЕННОСТИ СУЩЕСТВУЮЩИХ РЕШЕНИЙ

Краткое описание дополненной нотации Бэкуса Наура и контекстно-свободной грамматики

Согласно определениям [4, с. 105, 106] и [4, с. 111], контекстно-свободной грамматикой называется четвёрка $G = (T, N, P, S)$, где

- (1) N – конечное множество *нетерминальных символов*, или *нетерминалов* (иногда называемые синтаксическими переменными);
- (2) T – не пересекающееся с N конечное множество *терминальных символов*, или *терминалов* (иногда называемых *словами* или *токенами*);
- (3) P – конечное подмножество множества

$$N \times (N \cup T)^*$$

(элемент (a, b) множества P называется *правилом* (или *продукцией*) и записывается в виде: $a \rightarrow b$);

- (4) S – выделенный символ из N , называемый *начальным* (или *стартовым*) *символом*.

Здесь и далее правило вида $A \rightarrow \epsilon$ будет означать правило пустой строки, если не указано другое.

Дополненная нотация Бэкуса Наура определяет правила следующим образом: $Rule = elements\ crlf$, где:

- $Rule$ – имя правила (лексического или синтаксического), представленное в виде последовательности символов. Имя правила не чувствительно к регистру, т. е. имена Rule, rulE, RULE определяют одно и то же правило;
- $=$ – метасимвол, отделяющий имя правила от его тела;

- *elements* – непустая последовательность грамматических символов, т. е. имён правил или *терминальных сущностей*. *Терминальные сущности* представлены в виде последовательности символов, заключённых в двойные кавычки, которые кодируются в неотрицательные целые числа. Следовательно, данную последовательность можно записать в виде последовательности целых положительных чисел, разделённых между собой точкой («.») либо одиночным пробелом. Для указания основания системы счисления каждый код символа предваряется одним из перечисленных ниже префиксов:
 - *%d* – десятичное основание (10);
 - *%b* – двоичное основание (2);
 - *%x* – шестнадцатеричное основание (16);
- *crlf* – символ конца строки (символ возврата каретки, следующий за переводом строки).

Каждое правило может определять либо продукцию для нетерминала в контекстно-свободной грамматике, либо шаблон (последовательность символов) для терминального символа. Правило для стартового символа *S* определяется в первую очередь. Имя правила, правая часть которого содержит только терминальные сущности, является терминальной сущностью. Множество всех терминальных сущностей составляет множество терминалов грамматики *T*. Множество правил, не являющихся терминальными сущностями, составляет множество продукции грамматики (их имена составляют множество нетерминалов грамматики *N*). Имя первого определённого правила является стартовым символом грамматики.

Нотация дополнительно определяет несколько операторов:

- “*m*nR*” – оператор повторения, где *m* и *n* – целые положительные числа. Он предшествует имени грамматического символа *R* (имени правила или терминальной сущности). Означает: “*R* следует не менее *m* раз и не более *n* раз”. Например: запись «*1*3+*» означает – следует “+” или “++” или “+++” (т. е. 1, 2 или 3 раза следует символ “+”). Если *m* не указано, то оно равно 0. Если *n* не указано, то оно равно бесконечности. Таким образом, запись «**1R*» озна-

чает 0 или 1 раз идёт символ R, а запись «1*R» означает 1 раз или более идёт символ R;

- Группирующие круглые скобки “(”, “)” – используются для группировки грамматических символов в сочетании с оператором повторения. Оператор повторения предшествует открывающейся скобке. Для того чтобы использовать скобки в качестве терминальных символов, достаточно заключить их в двойные кавычки;
- Квадратные скобки “[”, “]” – используются для указания выборочной последовательности грамматических символов. Это не что иное, как сокращение выражения “*1()”, т. е. вместо записи «*1(foo bar)» можно использовать «[foo bar]»;
- “/” – оператор выбора. Используется между последовательностями элементов одного и того же правила для обозначения альтернативной последовательности (которая следует справа от символа “/”). Может применяться многократно. Например, правило вида $abc = "a" / "b" / "c"$ означает либо “a”, либо “b”, либо “c”.

Анализ входных форматов грамматики существующих генераторов синтаксического разбора

Для анализа формата описания грамматики рассмотрим несколько известных приложений генерации синтаксических анализаторов с открытым исходным кодом:

- Parglare – приложение с открытым исходным кодом, реализующее генератор синтаксических анализаторов типа GLR и LR [5]. Написано на языке программирования Python;
- ANTLR v4 – генератор нисходящих анализаторов (типа LL) для формальных языков [6]. ANTLR преобразует контекстно-свободную грамматику в виде расширенной нотации Бэкуса Наура в программу на C++, Java, C#, JavaScript, Go, Swift, Python. Реализован на языке Java;
- GNU Bison – программа, создающая синтаксические анализаторы по переданному ей описанию грамматики [7]. Реализована на языке C.

Все перечисленные программы берут за основу синтаксиса описания грамматик форму Бэкуса Наура. Однако каждая программа имеет свои собственные интерфейс и функционал для работы с описанием грамматики. Кроме того, каждая программа по-своему задаёт структуру входного файла грамматики, имеет свои расширения для формы Бэкуса Наура. В следующих подразделах перечислены основные элементы синтаксиса описания грамматики для каждой указанной программы. Для сравнения синтаксиса описания грамматики использована простая грамматика арифметических выражений, которая описана ниже в формате ABNF [1]. Терминальные сущности, представленные в виде литералов строк, выделены красным цветом.

$S = 1*(E \text{ newline})$

$E = E \text{ "+" } T / E \text{ "-" } T / T$

$T = T \text{ "*" } F / T \text{ "/" } F / F$

$F = \text{"(" } E \text{ ")" } / \text{number}$

$\text{number} = 1*\text{digit} [\text{"." } 1*\text{digit}]$

$\text{digit} = \text{"0"} / \text{"1"} / \text{"2"} / \text{"3"} / \text{"4"} / \text{"5"} / \text{"6"} / \text{"7"} / \text{"8"} / \text{"9"}$

$\text{newline} = \text{\%x2424}$

Листинг 1. Грамматика арифметических выражений в нотации ABNF.

Синтаксис описания грамматики, предлагаемый программой Parglare

Входной файл описания грамматики имеет расширение *.pg* и состоит из следующих секций, расположенных в указанном ниже порядке:

- Секция инструкций импорта. Она не является обязательной;
- Секция правил, содержащая одно или множество правил грамматики;
- Секция терминалов грамматики "terminals". Имя секции (terminals) указывается обязательно.

Инструкции импорта принимают имя файла грамматики и позволяют сослаться на его содержимое по имени импортируемой грамматики. Именем грамматики считается по умолчанию имя файла, в котором она написана. Таким образом инструкции импорта позволяют определять одну грамматику в не-

скольких файлах по частям. Синтаксис инструкции выглядит следующим образом:

- `import './.../myGrammarPart1.pg';`

Правило грамматики имеет следующий вид:

- *ruleName: expression ;*

Здесь *ruleName* означает имя нетерминального символа грамматики (синтаксической переменной), а *expression* – последовательность имён нетерминальных или терминальных символов грамматики, разделённых пробелом. Каждое правило заканчивается точкой с запятой (“;”). Терминальным символом может быть либо заранее предопределённое имя терминала, либо строка символов, заключённая в двойные кавычки.

Имена терминалов определяются в секции `terminals`, следующей за всеми правилами грамматики. Определение терминала выглядит следующим образом:

- `termName: stringOrRegex ;`

Здесь `termName` – имя терминала, а `stringOrRegex` – либо строка, заключённая в двойные кавычки, либо литерал регулярного выражения на языке Python (шаблон токена), либо пустая строка. Аналогично правилам грамматики каждое определение терминала заканчивается точкой с запятой. Имя `EMPTY` зарезервировано и используется как правило, обозначающее пустую строку. `Parlage` позволяет добавлять собственные лексические распознаватели к терминалам, у которых в правой части (после двоеточия) ничего нет кроме пробелов и одиночного символа двоеточия.

Согласно дополненной нотации Бэкуса Наура [1] каждое правило (синтаксическое и лексическое) должно заканчиваться символом конца строки. В `Parlage` вместо него используется точка с запятой. `Parlage` определяет собственный набор операторов, отличный от нотации ABNF. `Parlage` позволяет повторять определённый грамматический символ (терминал и нетерминал) один или более раз с помощью оператора “+”, который должен следовать сразу за именем грамматической сущности, в отличие от нотации, где оператор повторения

определяется между двумя числами, задающими диапазон повторений, перед именем грамматического символа. Символ “?” используется аналогично, но означает повторение 0 или 1 раз. Его можно использовать вместо EMPTY для выборочной последовательности грамматических символов. В качестве оператора выбора выступает символ вертикальной черты “|”. В файле грамматики можно оставлять однострочные комментарии в стиле языка C. Ниже дан пример грамматики арифметических выражений в форме синтаксиса Parglare.

```
S: E newline SE;  
SE: E newline SE | EMPTY;  
E: E "+" T | E "-" T | T;  
T: T "*" F | T "/" F | F;  
F: "(" E ")" | number;
```

terminals

```
number: \d+(\.\d+)?/;
```

```
newline: \n/;
```

Листинг 2. Грамматика арифметических выражений в нотации Parglare.

Большим преимуществом данной программы являются модульность грамматик, а также наличие отдельных секций для лексических и синтаксических правил, определяющих продукции и терминалы. Однако синтаксис регулярных выражений был взят из языка программирования Python, и их обработка выполняется на нём, т. е. регулярные выражения привязаны к реализации программы генератора, что является существенным недостатком данного приложения. Вторым недостатком является отсутствие поддержки группировки, определённой в нотации ABNF. Приложение не умеет группировать цепочку грамматических символов. Пара круглых скобок “(” и “)” используются лишь для группировки регулярных выражений, а не грамматических единиц.

Синтаксис описания грамматики, предлагаемый программой ANTLR 4

Файл грамматики имеет расширение .g4 и состоит из следующих разделов:

- Заголовок имени грамматики. Состоит из ключевого слова *grammar* и следующим за ним идентификатора. Заканчивается символом точка с запятой, следующим за идентификатором. Имя грамматики и имя файла должны совпадать;
- Раздел опций (*options {...}*). Данный раздел не является обязательным;
- Раздел токенов (*tokens {...}*). В данном разделе перечисляются имена токенов через запятую. Токены являются терминальными символами грамматики, для которых не определены лексические правила. Имена токенов прописные. Данный раздел не является обязательным;
- Раздел инструкций импорта. Не является обязательным;
- Раздел синтаксически управляемых действий (*actions {...}*). Содержит исполняемый Java-код. Не является обязательным;
- Раздел правил (синтаксических и лексических). Имя для лексического правила набирается в прописном регистре. Имя синтаксического правила должно начинаться со строчной буквы.

Правила грамматики могут быть представлены в следующих видах:

LEXRULE : lexems ;

syntaxRule : elements ;

Здесь *LEXRULE* – имя лексического правила, задающего шаблон для терминала (с именем *LEXRULE*), *lexems* – элементы лексического правила, разделённые пробелами между собой. Элементами лексического правила могут быть последовательности символов, заключённых в одинарных кавычках, имена других лексических правил, или *фрагменты*. Фрагмент – это часть грамматики, которая может быть только частью лексического правила. Фрагменты нельзя использовать в синтаксических правилах. Лексическое правило может стать фрагментом. Для этого необходимо предварить имя правила ключевым словом *fragment*. Ключевое слово указывается перед именем правила отдельно.

Имя *syntaxRule* обозначает продукцию в грамматике с одноимённым не-терминалом в качестве заголовка и *elements* в качестве тела. Здесь *elements* обозначает последовательность имён правил (любых, как лексических так и синтаксических) и строк символов, заключённых в одинарные кавычки.

Каждое правило заканчивается символом «точка с запятой» и отделяется от другого правила символом новой строки. Первое определённое правило является продукцией для стартового символа грамматики. ANTLR также позволяет группировать грамматические элементы в цепочки, как в нотации ABNF. ANTLR предоставляет собственные операторы повторения, перечисленные ниже:

- (1) “+” – оператор повторения 1 или более раз;
- (2) “*” – оператор повторения 0 или более раз;
- (3) “?” – оператор повторения 0 или 1 раз. (аналогично Parglare).

Каждый из операторов, перечисленных выше, повторяет цепочку грамматических символов, расположенную слева от него. Оператор выбора реализован так же, как в программе Parglare. Ниже описана грамматика арифметических выражений в форме синтаксиса ANTLR.

```
grammar Expressions;
prog: (expr NEWLINE)+ ;
expr: expr ( '*' | '/' ) term | term ;
term: term ( '+' | '-' ) factor | factor ;
factor: NUMBER | '(' expr ')' ;
NEWLINE : [\n];
NUMBER : [0-9]+(\.[0-9]+)? ;
```

Листинг 3. Грамматика арифметических выражений в нотации ANTLR.

Формат описания грамматики, предлагаемый ANTLR, обладает большим недостатком. В первую очередь, от регистра имени зависит тип правила (определение терминала, или синтаксической продукции) для грамматики. Зависимость от регистра может внести путаницу при создании грамматики. Кроме того, нотация ABNF явно указывает, что имена правил не зависят от регистра букв.

Лучшим решением является определение отдельной секции для лексических правил, как это было сделано в приложении Parglare. Также программа ANTLR не умеет разрешать неявную левую рекурсию, следовательно, она перекладывает устранение левой рекурсии из грамматики на плечи рядового пользователя, в то время как существует алгоритм устранения левой рекурсии из [5, с. 180], позволяющий решить данную проблему.

Достоинством ANTLR является поддержка круглых скобок, группирующих цепочки грамматических символов. Благодаря им можно сократить дублирование, возникающее при перечислении альтернатив синтаксических правил, у которых есть общие части цепочек.

Синтаксис описания грамматики, предлагаемый программой GNU Bison

Входной файл грамматики для приложения GNU Bison имеет имя, оканчивающееся на `.y`, оно состоит из четырёх секций, расположенных в порядке, указанном ниже:

- Секция пролога. Содержит макроопределения, макросы, объявления переменных и функций на языке C. Может отсутствовать;
- Секция определений. В данной секции определяются имена терминальных и нетерминальных символов грамматики. Для объявления терминала необходимо объявить токен с его именем. Имя токена должно быть полностью в верхнем регистре. Токены объявляются инструкцией `%token`. Секция не является обязательной;
- Секция правил. Содержит непустую последовательность определений правил трансляций. Данная секция начинается с заголовка, состоящего из двух символов: `'%%'`. Далее, после заголовка, с новой строки идут определения продукций грамматики. Синтаксис определения аналогичен синтаксису приложения Parglare, но с ограничением на имена нетерминалов – они не должны состоять из одних прописных букв, т. к. все имена токенов должны быть полностью в верхнем регистре. Пустое тело правила (обозначающее пустую строку) обозначается как `%empty`;

- Секция эпилога. Данная секция следует после секции правил. Для её отделения от секции правил используется тот же заголовок ‘%%’. Не является обязательной;

Ниже описана грамматика арифметических выражений в виде файла грамматики GNU Bison:

```
%token NUMBER

%%

s: expr line se;
se: expr line se | %empty;
line: expr '\n';
expr: expr '+' term | expr '-' term | term;
term: term '*' factor | term '/' factor | factor;
factor: '(' expr ')' | NUMBER;

%%

yylex(){
return getNumber();
}
```

Листинг 4. Грамматика арифметических выражений в нотации GNU Bison.

Одним из главных недостатков данного формата является отсутствие задания лексических правил для терминальных символов грамматики. Всю работу по сборке последовательности символов в терминалы (в слова) приходится выполнять вручную, с использованием предопределённой функции *yylex*. Она должна вернуть положительное целое число, связанное с определённым токеном. Программа автоматически назначает определённым токенам их коды, так что для возвращения кода достаточно набрать имя токена. Функция также должна сохранить распознанную лексему в переменной *yylval*. В листинге 4 работа функции переложена на функцию *getNumber*, которая собирает число (как целое, так и дробное) в переменную *yylval* и возвращает код токена NUMBER. Ещё одним

существенным минусом является отсутствие поддержки операторов группировки и повторения, определяемых ABNF.

Также программа использует имена токенов (терминалов) в верхнем регистре, что опять может привести к ошибкам в определении имён правил грамматики.

Однако программа способна справляться с леворекурсивными грамматиками, в отличие от приложения ANTLR. Она также способна обрабатывать правую рекурсию. Программа способна работать с генератором лексических анализаторов Lex, позволяя ей определять лексические правила для токенов, задаваемых текущим форматом грамматики.

ФОРМАТ ДАННЫХ JSON ДЛЯ КОНТЕКСТНО-СВОБОДНЫХ ГРАММАТИК

Во всех перечисленных форматах есть определённые особенности, присущие каждому. Однако они не определяют все компоненты контекстно-свободной грамматики в явном виде. В каждом из них используются свои собственные правила, диктующие, какие элементы относятся к определённому компоненту грамматики. Ниже представлен новый формат описания грамматик, лишённый данного недостатка. В нём для каждого компонента определена соответствующая секция.

Контекстно-свободная грамматика состоит из четырёх элементов: множества терминалов T , нетерминалов N , продукций P и стартового символа S . Определим для неё JSON-документ, который будет хранить информацию об этих множествах. JSON-объект будет состоять из четырех свойств:

- Свойство “terms” представляет множество терминалов грамматики в виде JSON-объекта, содержащего коллекцию пар *ключ–значение*. *Ключом* является имя терминала, которое должно быть уникальным в пределах объекта. *Значением* является строка, содержащее регулярное выражение. Синтаксис регулярного выражения был взят из [4, с. 124] и дополнен в следующем параграфе настоящей статьи;
- Свойство “nonTerms” представляет JSON-массив, состоящий из строк, обозначающих имена нетерминалов грамматики;

- Свойство “start” представляет собой строку, обозначающую имя нетерминала из массива “nonTerms”, который является стартовым символом грамматики;
- Свойство “productions” хранит продукции грамматики в JSON-массиве. Элементами массива являются JSON-объекты с единственным свойством. Имя такого свойства должно присутствовать в массиве nonTerms (т. е. быть нетерминалом грамматики). Значением такого свойства являются либо массив грамматических символов, либо литерал null, либо одиночный грамматический символ, выраженный строкой. Все грамматические символы, составляющие тела продукций, должны быть выражены в качестве строк, обозначающих имена терминалов или не терминалов (т. е. либо являются элементами массива nonTerms, либо свойствами объекта terms). В случае, когда тело продукции состоит из более чем одного грамматического символа, необходимо использовать JSON-массив. В массиве перечисляются строки (имена терминалов или нетерминалов) в определённом порядке. Литерал null можно использовать для обозначения тела продукции, порождающей пустые строки, вместо имени терминала, значение которого в объекте terms равно null. Для создания альтернатив к определённому правилу с именем N необходимо создать новый JSON-объект в данном массиве со свойством, имеющим такое же имя N.

Во многих языках программирования зарезервированы специальные ключевые слова. Можно определить дополнительное, необязательное свойство в объекте грамматики с именем *keywords*, значением которого будет массив строк, представляющих ключевые слова.

Дополнения к регулярным выражениям

Регулярные выражения снабжены дополнительными операциями:

- “+” – положительное замыкание Клини. Означает повторение шаблона 1 или более раз;
- [a-z] – классы символов. Являются сокращением многократного использования оператора объединения “|”. Например, указанная запись «[a-z]» означает «(a|b|c|...|z)»;

- “_” – зарезервированное регулярное выражение, которое обозначает любую строку из одного символа, т. е. любой одиночный символ;
- “@” – экранирующий символ. Любой одиночный символ, который следует за ним, воспринимается буквально (как операнд). Чтобы написать выражение для строки из символа “@”, достаточно его продублировать – “@@”. Во многих языках программирования в синтаксисе регулярных выражений для аналогичных действий используют символ обратного следа “\”. Но поскольку стандартом [3] определено его использование при разборе JSON-документов, в целях совместимости со стандартом для регулярных выражений был определён экранирующий символ собаки “@”;
- JSON-литерал null обозначает регулярное выражение для множества пустых строк.

При использовании последовательности числовых кодов символов налагается ещё одно ограничение: числовые коды со значениями 0 и 1 зарезервированы для символа “_” и символа пустой строки соответственно. При их указании в последовательности нужно понимать, что в данном случае символы будут интерпретироваться как “_” (в случае 0) или как маркер пустой строки (в случае 1).

Таким образом, опираясь на определённый JSON формат грамматики из предыдущего параграфа и синтаксиса регулярных выражений с вышеперечисленными дополнениями и ограничениями грамматика арифметических выражений представлена следующим образом:

```
{
  "terms" : {
    "newline" : "\n",
    "empty" : null,
    "number" : "[0-9]+.[0-9]+(((E|e)(-|empty)[0-9]+)|empty)",
    "+" : "+",
    "-" : "-",
    "*" : "*",
    "/" : "/",
    "(" : "(",
    ")" : ")"
  },
  "nonTerms" : ["S", "SE", "E", "T", "F"],
  "productions" : [
    {"S" : ["E", "newline", "SE"] },
    {"SE" : ["E", "newline", "SE"] }, {"SE" : "empty" },
    {"E" : ["E", "+", "T"] }, {"E" : ["E", "-", "T"]}, {"E" : "T"},
    {"T" : ["T", "*", "F"] }, {"T" : ["T", "/", "F"]}, {"T" : "F"},
    {"F" : ["(", "E", ")"] }, {"F" : "number"}
  ],
  "start" : "S"
}
```

Листинг 5. Грамматика арифметических выражений в нотации JSON.

ГЕНЕРАЦИЯ ПРОГРАММЫ СИНТАКСИЧЕСКОГО РАЗБОРА ИЗ JSON-ДОКУМЕНТА КОНТЕКСТНО-СВОБОДНОЙ ГРАММАТИКИ

Опишем концепцию генератора синтаксических анализаторов, который будет использовать предложенный JSON-формат.

Описание предметной области

Для составления схемы построения и работы генератора необходимо выделить сущности и их обязанности. Для этого была создана карта сущностей (рис. 1) и на её основе составлен список обязанностей и основных функции, требуемых при создании приложений генераторов разборщиков.

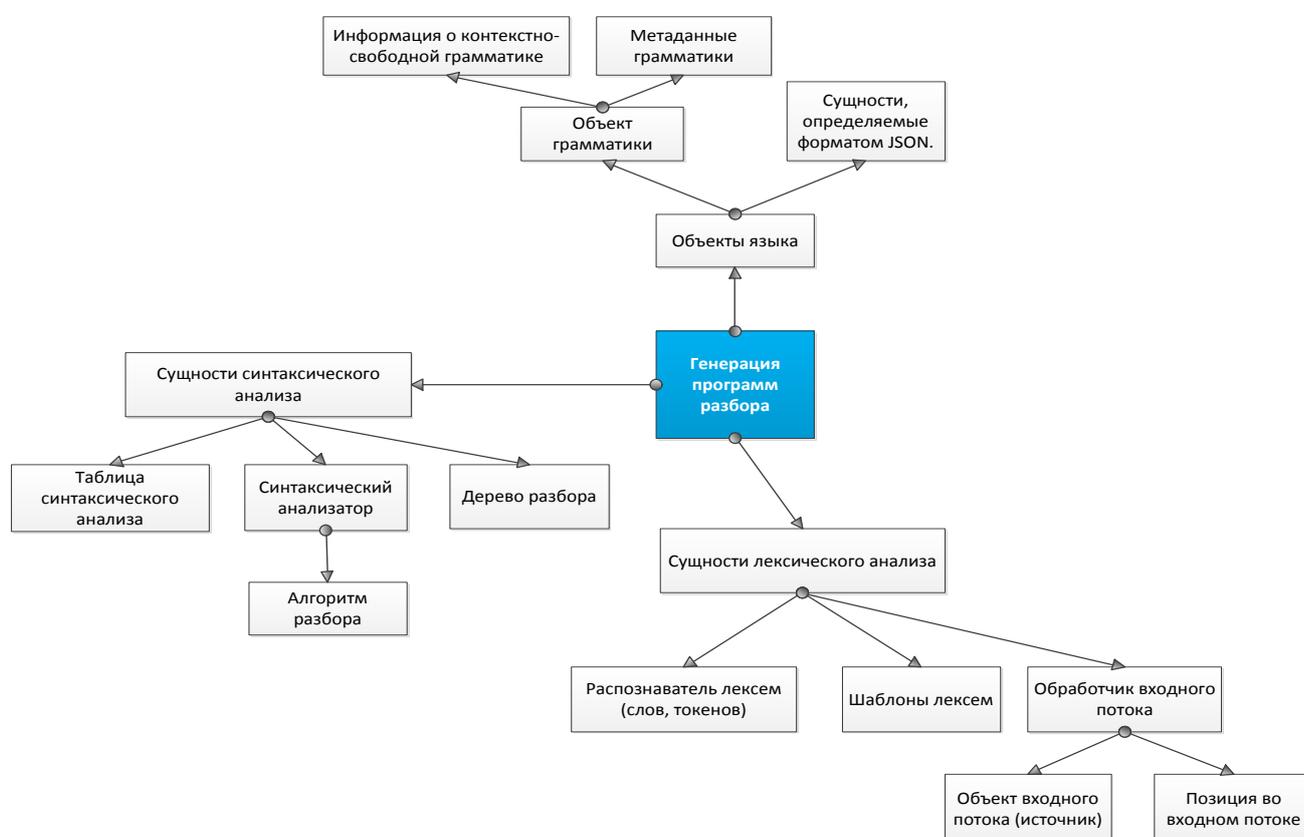


Рисунок 1. Карта основных сущностей генератора программ разбора.

Перечислим основные обязанности приложения:

- Обработка JSON-документов и их представление в виде дерева;
- Создание объекта грамматики:
 - Обработка дерева JSON-документа, хранящего информацию о грамматике;

- Проверка JSON-документа на соответствие формату, описанному в данной статье;
- Выделение дополнительной информации, не относящейся к определению контекстно-свободных грамматик, и её сохранение в виде отдельного объекта;
- Преобразование грамматики к виду, пригодному для применения одного из алгоритмов разбора:
 - Устранение левой рекурсии;
 - Устранение пустых правил. Если язык, задаваемый грамматикой, порождает пустые строки, то в грамматике должно присутствовать правило вида $S \rightarrow \epsilon$, где нетерминал S не встречается в правых частях любых других правил грамматики;
 - Устранение циклов;
- Создание лексического анализатора по шаблонам терминалов грамматики:
 - Чтение из источника (входного потока);
 - Идентификация распознанной лексемы;
 - Оповещение о лексической ошибке в случае невозможности распознать лексему;
- Генерация таблицы синтаксического анализа по одному из алгоритмов, задаваемых пользователями (LL, LR, LALR, SRL, GLR);
- Создание объекта разборщика на основе таблицы синтаксического анализа и лексического анализатора;
- Анализ текстовых файлов построенным синтаксическим анализатором:
 - Выявление синтаксических и лексических ошибок;
 - Построение дерева разбора, в случае отсутствия ошибок;
 - Восстановление при обнаружении ошибок.

Структура генератора программ разбора

Основываясь на карте сущностей и списке обязанностей, представленных в предыдущем параграфе, можно построить набор компонентов и модулей генератора.

Основные структурные модули приложения:

- `JsonParser` – основной класс, объединяющий процедуры разбора JSON документов. Имеет два метода: *parseFile* и *parseStream*, которые принимают в качестве параметра либо имя файла, либо поток целочисленных, положительных символьных кодов;
- `JsonElement` – корень иерархии типов, задаваемых стандартом JSON [3]. С помощью него извлекаются значения данных различных типов;
- `JsonObject` – тип объекта, представляющий структуру и основное содержимое прочитанного JSON-документа. Предоставляет два метода: *getProperty* и *getElement*. Оба метода принимают строку, обозначающую имя свойства. Метод *getProperty* ищет указанное свойство в самом `JsonObject` без учёта вложенности. Метод *getElement* учитывает вложенные объекты. Оба метода возвращают экземпляр класса `JsonElement`, который хранит значение данного свойства;
- `ILexer` – интерфейс взаимодействия с лексическим анализатором, позволяющий получать токены по одному. Определяет единственный метод *recognize*, который принимает в качестве параметра указатель входного потока. `ILexer` читает лексемы посимвольно с помощью `ILexerIO`, предоставляя ему всю работу с входным потоком;
- `ILexerIO` – интерфейс взаимодействия с входным потоком, откуда лексический анализатор читает символы и собирает их в лексемы. В нём определены следующие методы:
 - *ungetch(char c)*: сохраняет символ *c* в буфер;
 - *getch()* – вернёт символ с буфера, если он не пуст, иначе вернёт текущий символ с входного потока. После этого переместит указатель входного потока на один символ вперёд;

- `getInput()` – возвращает текущий символ входного потока. После этого переместит указатель входного потока на один символ вперёд;
- `Token` – тип объекта, содержащего информацию о распознанной лексеме, выраженную в трёх полях:
 - `name` – имя символа (распознанной лексемы);
 - `value` – сама лексема (прочитанная строка символов);
 - `type` – тип символа. Определяет, является ли символ терминалом, нетерминалом, символом конца потока или ошибки;
- `DFA Lexer` – реализация интерфейса `ILexer`. Основана на архитектуре лексического анализатора программы `Lex` и детерминированного конечного автомата, который строится по алгоритмам [8, с. 206] и [8, с. 213]. Реализует модуль лексического анализа;
- `App` – основной класс приложения, инкапсулирующий процесс построения анализатора и синтаксического разбора передаваемых файлов;
- `Grammar` – тип объекта, представляющий информацию о контекстно-свободных грамматиках. Может включать в себя объект типа `GrammarInfo`;
- `GrammarInfo` – тип объекта, содержащий дополнительные данные о грамматике (ключевые слова, терминалы, обозначающие комментарии, т. е. слова, которые надо игнорировать синтаксическому анализатору);
- `Parser` – абстрактный класс, представляющий синтаксический анализатор. Хранит два объекта: экземпляр класса `Grammar` и `DFA Lexer`. Его подклассы должны реализовать один из алгоритмов синтаксического анализа;
- `ParsedTree` – дерево разбора, создаваемое объектом типа `Parser`. Содержит информацию о синтаксической структуре разобранного текста.

Схема работы генератора

Главным классом приложения является *App*. Он отвечает за работу остальных модулей программы. Принимает на вход от пользователя имя файла, содержащее описание грамматики в формате JSON. Далее, при успешном построении анализатора ожидает на вход текстовые файлы для синтаксического разбора. Схема работа данного класса представлена на диаграмме последовательности UML (рис. 2).

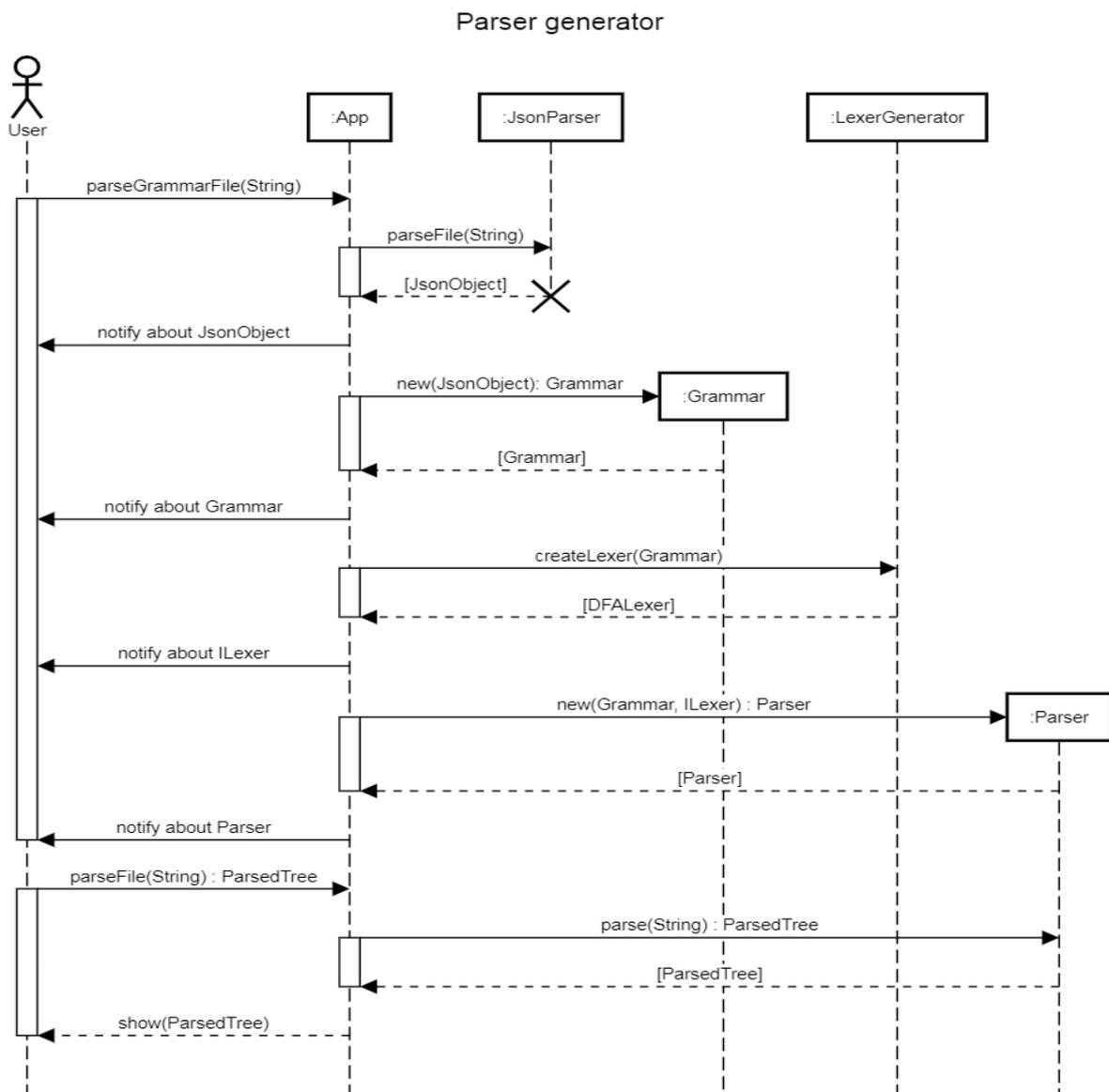


Рисунок 2. Диаграмма последовательности UML для программы генератора синтаксических анализаторов.

ЗАКЛЮЧЕНИЕ

Предложенный формат описания грамматик позволяет упростить процесс создания программы разбора. С помощью формата данных JSON и отдельных секций для каждого компонента грамматики очень легко и просто отделить лексические правила и единицы от синтаксических. Использование данного формата с предложенной архитектурой генераторов синтаксического разбора позволяет сконцентрироваться на разработке контекстно-свободного языка, а разделение обязанностей по чтению и распознаванию лексем в два отдельных модуля предоставляет возможность разработчику выбирать тип источника, который можно преобразовать в последовательность символов. Также данный формат позволяет легко расширять или модифицировать язык, определяемый грамматикой.

СПИСОК ЛИТЕРАТУРЫ

1. Standard RFC 5234: Augmented BNF for Syntax Specifications: ABNF. URL: <https://tools.ietf.org/html/rfc5234#section-2.1>
 2. Standard RFC 8259: The JavaScript Object Notation (JSON) Data Interchange Format. URL: <https://tools.ietf.org/html/rfc8259>
 3. Standard ECMA 404: The JSON Data Interchange Standard. URL: <https://www.json.org/json-en.html>, <https://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf>
 4. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции: В 2-х томах. Т. 1: Синтаксический анализ. М.: Мир, 1978.
 5. Parlage // Github. URL: <https://github.com/igordejanovic/parglare>
 6. ANTLR: ANOther Tool for Language Recognition. URL: <https://www.antlr.org/>
 7. GNU Bison: Manual. URL: <https://www.gnu.org/software/bison/manual/bison.html>
 8. Ахо А.В., Лам М.С., Сети Р., Ульман Дж. Д. Компиляторы: принципы, технологии и инструментарий, 2-е изд.: Пер. с англ. М.: Вильямс, 2008. 1184 с.
-

DESCRIPTION OF CONTEXT FREE GRAMMARS IN JSON FORMAT FOR PARSER GENERATORS

O. K. Osipov ¹

Bauman Moscow State Technical University

¹osipovok@student.bmstu.ru

Abstract

Analysis of various presentations for context free grammars provided with parser generators. A new description format of context free grammars is proposed. Given a representation of context free grammar in JSON format. The concept of a new parser generator based on JSON data format of describing context free grammars is presented. Described a parser generation scheme based on that concept.

Keywords: *JSON-document, context free grammars, lexeme, Backus Naur Form, parsing tree, terminal symbols (tokens), deterministic finite state automata, parser, Parglare, ANTLR.*

REFERENCES

1. Standard RFC 5234: Augmented BNF for Syntax Specifications: ABNF. URL: <https://tools.ietf.org/html/rfc5234#section-2.1>
2. Standard RFC 8259: The JavaScript Object Notation (JSON) Data Interchange Format. URL: <https://tools.ietf.org/html/rfc8259>
3. Standard ECMA 404: The JSON Data Interchange Standard. URL: <https://www.json.org/json-en.html>, <https://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf>
4. *Alfred V. Aho, Jeffrey D. Ullman.* The theory of parsing, translation and compiling: In 2 volumes. V.1: Parsing. Moscow, Mir Publ., 1978.
5. Parglare // Github. URL: <https://github.com/igordejanovic/parglare>
6. ANTLR: ANOther Tool for Language Recognition. URL: <https://www.antlr.org/>
7. GNU Bison: Manual. URL: <https://www.gnu.org/software/bison/manual/bison.html>

8. *Alfred V. Aho, Monica S. Lam, Ravi Sethi, Jeffrey D. Ullman. Compilers: Principles, Techniques and Tools, 2nd ed. Moscow, Addison-Wesley Publ., 2008. 1184 p. (In Russian).*

СВЕДЕНИЯ ОБ АВТОРЕ



ОСИПОВ Олег Константинович – студент магистратуры факультета Информатики и систем управления Московского государственного технического университета им. Н.Э. Баумана, направление подготовки – «Программная инженерия».

Oleg Konstantinovich OSIPOV, student of 1 grade of master program of Computer Science and Control Systems department of Bauman Moscow State Technical University.

Email: osipovok@student.bmstu.ru

Материал поступил в редакцию 15 августа 2020 года

УДК 004.550

ЦИФРОВОЙ РЕПОЗИТОРИЙ "GEOLOGYSCIENCE.RU": ОТКРЫТЫЙ ДОСТУП К НАУЧНЫМ ПУБЛИКАЦИЯМ ПО ГЕОЛОГИИ РОССИИ

М. И. Патук¹, В. В. Наумова², В. С. Ерёменко³

Федеральное государственное бюджетное учреждение науки «Государственный геологический музей им. В.И. Вернадского РАН»

¹patuk@mail.ru, ²naumova_new@mail.ru, ³vitaer@gmail.com

Аннотация

Описаны новые подходы, связанные со сбором данных из разнородных информационных систем доступа к научным публикациям с использованием открытых международных стандартов и протоколов для формирования систем открытого доступа к научным геологическим публикациям. На основе разработанных и адаптированных подходов и технологических решений реализован комплекс программ информационно-аналитической системы доступа к научным публикациям, реализующей функции сбора, поиска, каталогизации, фильтрации и управления научными публикациями и их метаданными.

Ключевые слова: *информационные технологии, науки о Земле, репозиторий, научные публикации.*

Введение

Термин «открытый доступ» впервые был упомянут на Будапештской конференции по открытому доступу в феврале 2002 г. [1]. С тех пор его смысл практически не изменился: Open Access определяется как бесплатный (free), оперативный (immediate), постоянный (permanent), полнотекстовый (fulltext), онлайн-овый (online) доступ к научным публикациям.

Открытый доступ — это бесплатный доступ читателей к онлайн-овым научным публикациям с правом читать, загружать, копировать, распространять, печатать, искать, ссылаться на полнотекстовые статьи, индексировать и т. п., то есть использовать с любой законной целью без финансовых, юридических или технических препятствий.

Выделяют два основных технологических направления: журналы открытого доступа и архивы (репозитории) открытого доступа. Оба направления – способы научного общения. Журналы открытого доступа публикуют прореферированные статьи, а репозитории собирают документы – не обязательно прошедшие реферирование и не обязательно статьи. Репозиторий собирает «свои» работы, т. е. труды сотрудников данного учреждения, и этим принципиально отличается от библиотеки. Журналы открытого доступа и репозитории не являются взаимоисключающими – они дополняют друг друга.

Цифровые репозитории служат реальным показателем качества деятельности университета/института и показывают научную, социальную и экономическую значимость исследовательских работ и таким образом демонстрируют статус и общественное значение университета/института.

Для практической реализации моделей открытых архивов предполагается:

- обмениваться метаданными, а не самими цифровыми объектами;
- использовать асинхронную технологию сбора данных;
- сформировать две группы участников системы «Открытого архива»: поставщики данных и поставщики услуг.

Поставщики данных (открытые архивы, репозитории) обеспечивают свободный доступ к метаданным и бесплатный доступ и пользование ресурсами, а также простоту в работе, не требующую создания каких-либо специальных коллективов и поэтому открывающую двери для участия малых организаций,

Поставщики услуг используют интерфейсы открытых архивов поставщиков данных, собирают и хранят метаданные, выбирают некоторые специализированные коллекции от поставщиков данных, пополняют состав метаданных и обогащают метазаписи, обеспечивают обслуживание на основе метаданных.

Разработано и успешно работает бесплатное программное обеспечение с открытым кодом для создания и поддержки таких OAI-совместимых архивов: наиболее популярные E-prints (<https://www.eprints.org/>), Dspace (<http://dspace.org/>).

В ходе 69-й Генеральной конференции ИФЛА на семинаре «Информационные технологии и работа группы метаданных Dublin Core» были сформулированы принципы, на которых базируется идеология «Открытого архива»: консор-

лидация в мировом масштабе архивов научных материалов; свободный доступ к архивам (к метаданным); согласованные интерфейсы архивов и поставщиков информации; простота использования; применение существующих стандартов – HTTP, XML, Dublin Core, MARC, MARCXML [2].

В последнее десятилетие наблюдается качественно новый уровень в организации хранения и предоставления «открытых научных данных». Активно развиваются системы и платформы, которые обеспечивают весь процесс управления данными – от публикации автором до анализа и повторного использования этих данных любым исследователем или системой. В концептуальной основе новых систем лежат «принципы цитирования данных» DataCite [3], FAIR-принципы [4] и рекомендации мировых ассоциаций по обмену данными: The International Science Council (ISC) [5], The Research Data Alliance (RDA).

Современной формой хранения и предоставления научной информации стали наборы данных. Набор данных – это контейнер, содержащий данные, метаописание (в формате Dublin Core или DataCite) и уникальный идентификатор (например, DOI). Мировые центры данных предоставляют доступ к своим хранилищам по протоколам Open Archives Initiative (OAI). Организация научной информации в виде наборов данных и доступность их метаданных по протоколам OAI позволяет упростить автоматизацию процессов поиска геологической информации по России. Надежность данных подтверждается указанием авторства, выходных данных статьи, проектов, программ, в рамках которых проводились исследования.

В рамках разработки Информационно-аналитической геологической среды "GeologyScience.ru", которая обеспечивает единую точку доступа к геологическим данным и системам их обработки на территории России [6–8], создан и поддерживается блок управления научными публикациями – <http://repository.geologyscience.ru/> (Рис. 1).

Блок разрабатывается как самостоятельный проект, но обладает свойствами для интеграции его в территориально-распределенные системы: однородность данных, наличие базы метаданных в международных форматах, доступ через API, поддержка сквозной авторизации и разграничения прав, службы

мониторинга и статистики. На Рис. 2 показано, каким образом блок «Научные публикации» входит в Информационно-аналитическую геологическую Среду.

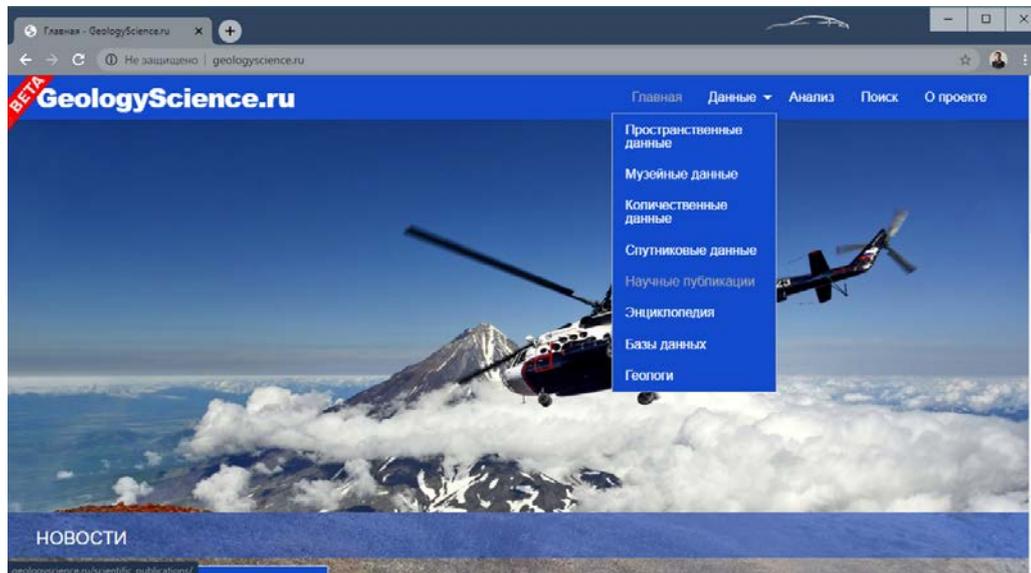


Рисунок 1. Главная страница портала “GeologyScience.ru”.

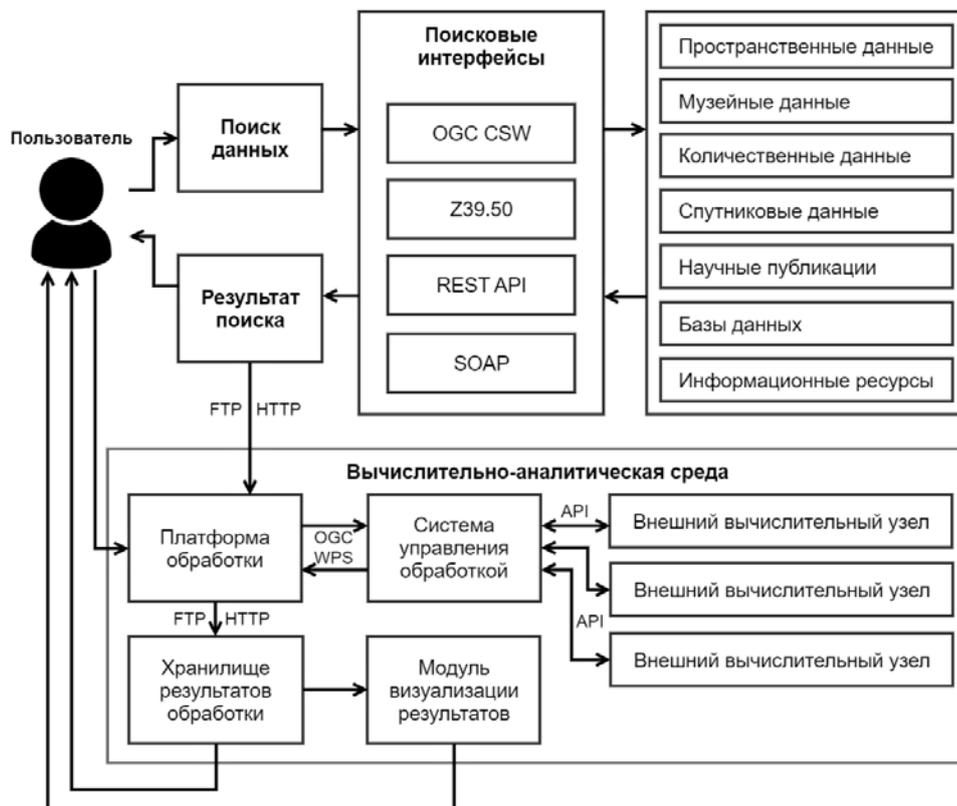


Рисунок 2. Обобщенная функциональная схема Информационно-аналитической геологической среды.

Авторами кратко сформулированы задачи блока «Научные публикации»:

- сбор и интеграция научных публикаций из территориально-распределенных источников;
- хранение данных метаданных в машиночитаемом виде и в принятых международных форматах;
- обеспечение доступности данных для пользователей и программ;
- тематическая адаптация интерфейса и функций системы.

Репозиторий «Геология России» создан на основе свободно распространяемого программного обеспечения DSpace 6.3. Системы подобного класса организованы для хранения и обмена цифровыми объектами и предоставляют доступ к информации по протоколу OAI. Имеются удобная система каталогов, наличие сервера сбора метаданных OAI-PMH, полнотекстовый поиск, основанный на поисковом инструменте Apache Lucene или Apache Solr, разграничение прав и поддержка протокола доступа LDAP, открытый код и большое сообщество пользователей и разработчиков по всему миру, возможность управлять и хранить цифровой материал любого типа.

Согласно общепринятому делению данный Репозиторий является тематическим. Тематика Репозитория – науки о Земле (геология, геохимия, петрология, минералогия, тектоника, геоморфология, вулканология, палеонтология, стратиграфия и т. п.). Основой информации для Репозитория являются научные статьи, монографии, диссертации, авторефераты диссертаций, тезисы докладов, материалы конференций, находящиеся в открытом доступе.

Коллекции из архивов и библиотечных каталогов могут быть доступны на интернет-ресурсе через три коммуникационных протокола: OAI-PMH (the Open Archives Initiative Protocol for Metadata Harvesting), Z39.50 или SRU (Search/Retrieve via URL).

Выбор коммуникационного протокола имеет большое влияние на функциональность, которую ресурс может предоставить конечному пользователю. Хотя все три протокола предоставляют стандарт для коммуникации между Репозиторием и библиотечными системами, коммуникационная парадигма, лежащая в основе, значительно различается. В то время как OAI-PMH позволяет ему собирать все записи метаданных из библиотек в центральный архив, Z39.50 и SRU

были разработаны для удаленного доступа и предоставления, поэтому записи метаданных остаются только у провайдера данных.

Сбор метаданных по протоколам Z39.50/SRU представляет значительную сложность. Протоколы Z39.50/SRU не предназначались для сбора метаданных, поэтому некоторая функциональность, требуемая для обеспечения эффективности и надежности процесса сбора, не была включена в проект протоколов.

В отличие от OAI-PMH, Z39.50-сервера доступны для значительного количества систем управления библиотеками, и его использование среди библиотек широко распространено. Большинство библиотек России в качестве системы управления библиотекой используется Ирбис (<http://irbis.elnit.org/>). Веб каталог системы, Web Ирбис, поддерживает сервер Z39.50, что позволяет попробовать осуществить сбор библиографических метаданных. В случае, когда сервер Z39.50 не настроен, можно воспользоваться функцией экспорта Web Ирбис, если она доступна.

Для поиска и извлечения информации из других репозиториев был создан скрипт на языке PHP. Обращение к сторонним репозиториям происходит по протоколу OAI_PMH.

Большинство репозиториев является институциональными, т. е. содержит информацию по многим направлениям, развиваемым в конкретном научном или образовательном учреждении, поддерживающем репозиторий.

Для улучшения поиска информации в репозитории к существующим стандартным в DSpace поисковым тегам был добавлен тег УДК (Универсальная Десятичная Классификация). Данная информация извлекается в полуавтоматическом режиме из выгруженного из DSpace бэкапа в текстовый файл с последующей загрузкой SQL-скриптом в таблицу PostgreSQL DSpace.

Данная информация позволяет в автоматическом режиме строить и обновлять предметный каталог репозитория, исходя из УДК русскоязычных публикаций по геологии России.

В текущий момент источниками данных для репозитория «Геология России» являются:

1. Репозитории:

- Репозиторий Томского политехнического университета. (<http://earchive.tpu.ru/>). Он является институциональным. Разделение по тематикам представлено слабо. Для отбора метаданных необходимо использовать словарь терминов. Платформа – DSpace.
- Репозиторий Института вулканологии и сейсмологии ДВО РАН. (<http://repo.kscnet.ru/>). Репозиторий тематический. Основная тематика связана с науками о Земле. В отборе метаданных по словарю терминов нет необходимости. Достаточно визуального контроля лог-файла импорта. Платформа – Eprints.
- Репозиторий Санкт-Петербургского государственного университета. (<https://dspace.spbu.ru>). Репозиторий институциональный. Четкое деление по тематикам, но незначительное количество данных. Отсюда взят раздел «Минералогия и кристаллография». Платформа – DSpace.

2. Электронные (цифровые) библиотеки:

- КиберЛенинка (<https://cyberleninka.ru/>). Открытая электронная библиотека. Четкое деление по тематикам. Присутствует раздел – Науки о Земле (earth-and-related-environmental-sciences). Требуется визуальный контроль лог-файла импорта из-за не всегда корректного отнесения публикаций к данному разделу. Данные можно брать из репозиториев по протоколу OAI-PMH. Но в отличие от репозиториев отсутствует сквозная нумерация данных, поэтому приходится сканировать данные по дате поступления.
- ELibrary (<https://www.elibrary.ru>). Самая представительная научная электронная библиотека. Но и самая требовательная к режиму доступа. В системе есть платное API для экспорта данных. Для реализации бесплатного доступа можно использовать свободно распространяемый скрипт – Subzer (<https://github.com/p1m-ortho/xs-sebzer>). Этот скрипт позволяет выгружать результаты поисковых запросов в формате BibTex. Система расширенного поиска в ELibrary достаточно функциональна, но, к сожалению, тематическим поиском пользоваться нельзя из-за

значительного количества некорректных результатов. Приходится пользоваться поиском по списку журналов и дате публикации.

3. PANGAEA. Сайт – <https://www.pangaea.de> – позиционирует себя как библиотека открытого доступа, предназначенная для хранения, публикации и распространения географически привязанных данных в науках о Земле. В основном представлены дополнительные данные (рисунки, фотографии, таблицы) к существующим публикациям. Как правило, имеется идентификатор DOI. В связи с тем, что представлены подробный предметный классификатор и ссылка на оригинал публикации, возможно использование этих данных для получения метаданных. Доступ к данным происходит по протоколу OAI-PMH. Данные сканируются по дате поступления.

МОДУЛИ ИМПОРТА ДАННЫХ

Для каждого из перечисленных источников создан свой модуль экспорта данных на языке PHP. Общая идеология работы схожа для всех модулей.

1. Сканирование источника:
 - сканирование по уникальным номерам (для репозиториев);
 - сканирование по дате поступления;
 - сканирование по списку предварительно отобранных ссылок.
2. Отбор данных:
 - выбор всех (если тематика – науки о Земле);
 - выбор по тематике (например – «earth-and-related-environmental-sciences»);
 - выбор по словарю терминов. Словарь терминов русскоязычный и содержит ~2100 терминов. Он создавался на основе ~2000 предварительно отобранных записей по интересующей нас тематике. Практика показала, что оптимальным оказалось наличие 3-х совпадений со словарем. При этом удается выбрать около 90 % источников, соответствующих тематике репозитория. Оставшиеся 10 % подвергаются ручной обработке. Эта информация служит основой для корректировки словаря.
 - для предотвращения дублирования импортируемых метаданных производится поиск наименования публикаций в базе данных DSpace – PostgreSQL. Дублированные метаданные исключаются из импорта.

3. Формирование XML-файлов со структурой DSpace Simple Archive Format с метаданными публикаций и экспорт файла публикации, при его наличии.
4. Импорт полученных данных в Репозиторий осуществляется стандартными средствами DSpace – импорт через простой архивный формат DSpace.

Анализ и обработка текстовых данных осуществляется в рамках Информационно-аналитической геологической среды "GeologyScience.ru", которая обеспечивает не только единую точку доступа к геологическим данным, но и к системам их обработки через Вычислительно-аналитическую геологическую среду <http://service.geologyscience.ru/> (Рис. 3) [9, 10].

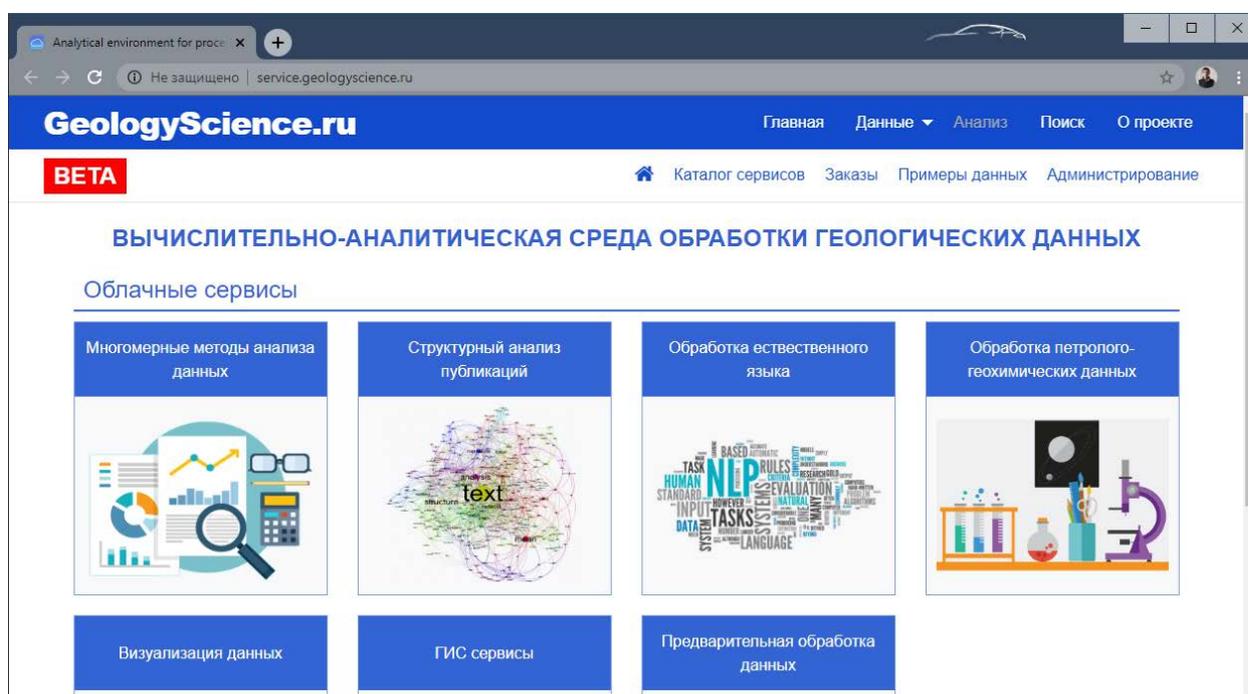


Рисунок 3. Единая точка входа к вычислительно-аналитической среде, входящей в состав "GeologyScience.ru".

В рассматриваемой вычислительно-аналитической среде для обеспечения единого подхода к вызову территориально-распределенных сервисов обработки и анализа данных используется стандарт OGC Web Processing Service (OGC WPS) в качестве промежуточного интерфейса.

Для анализа текстов в вычислительной среде доступны два удаленных вычислительных узла.

– **Структурный анализ публикаций.** В междисциплинарном центре математического и вычислительного моделирования (Университет Варшавы, Польша) разработан сервис для извлечения метаданных из научных публикаций [11]. Метаданные включают в себя авторов, аффилиацию, абстракт, ключевые слова, название журнала, объем, год выпуска, разобранные библиографические ссылки, структуру разделов документа, заголовки разделов и абзацы. Интерфейс взаимодействия с сервисами построен на основе REST-архитектуры.

– **Обработка естественного языка.** В Университете Шеффилда в рамках проекта GATE (General Architecture for Text Engineering) разработан ряд сервисов по обработке текстовых данных для различных языков [12]. Для обработки текстовых данных на русском языке предоставляются сервисы по определению частей речи слов, а также выделению именованных сущностей, таких, как имена и фамилии, названия организаций, географические названия, даты, денежные единицы и т. д. Интерфейс взаимодействия с сервисами построен на основе REST-архитектуры.

ЗАКЛЮЧЕНИЕ

Разработаны и адаптированы новые подходы, связанные со сбором разнородных данных из разнородных информационных систем доступа к научным публикациям с использованием открытых международных стандартов и протоколов в условиях системы ограничений, подразумевающей тематические, территориальные, временные ограничения, и методы, основанные на открытых международных стандартах метаданных научных публикаций и протоколов их обмена, теории построения распределенных информационных систем и поиска информации, для формирования тематических систем открытого доступа к научным публикациям.

Предложены новые технологические решения: сбора метаданных из разнотипных информационных систем, фильтрации собираемой информации с использованием УДК или специализированных тезаурусов, географического поиска на карте, организация тематического каталога.

На основе разработанных и адаптированных подходов и технологических решений реализован комплекс программ информационно-аналитической си-

стемы доступа к научным публикациям, реализующей функции сбора, поиска, категоризации, каталогизации, фильтрации и управления научными публикациями и их метаданными.

Благодарности

Научные исследования выполняются в рамках Государственного задания ФГБУН Государственного геологического музея им. В.И. Вернадского РАН по теме № 0140-2019-0005 «Разработка информационной среды интеграции данных естественнонаучных музеев и сервисов их обработки для наук о Земле».

СПИСОК ЛИТЕРАТУРЫ

1. Будапештская инициатива «Открытый доступ» / Budapest Open Access Initiative [Электронный ресурс]
URL: <https://www.budapestopenaccessinitiative.org> (дата обращения: 23.05.2020).
2. Технология открытых архивов / Open Archive Initiative, OAI [Электронный ресурс] URL: <https://openarchives.org/> (дата обращения: 23.05.2020).
3. *Brase J.* DataCite – A global registration agency for research data // In: Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology. IEEE, Beijing. 2009. P. 257–261. DOI: 10.1109/COINFO.2009.66
4. *Wilkinson M.D. et al.* The FAIR Guiding Principles for scientific data management and stewardship // In: Sci. Data 3:160018. 2016. P. 1.
DOI: 10.1038/sdata.2016.18.
5. *Claudia Emerson, Elaine M. Faustman, Mustapha Mokrane, Sandy Harrison.* World Data System (WDS) Data Sharing Principles // Zenodo. 2015.
<http://doi.org/10.5281/zenodo.34354>
6. *Naumova V.V., Belousov A.V.* Digital repository «Geology of the Russian Far East» – an open access to the spatially distributed online scientific publications // Russian Journal of Earth Sciences. 2014. Vol. 14, No. 1. P. 1–8.
7. *Наумова В.В., Платонов К.А., Еременко В.С., Патук М.И., Дьяков С.Е.* Информационно-аналитическая среда для поддержки научных исследований в геологии: текущее состояние и перспективы развития // Труды XVII Международной конференции «Распределенные информационно-вычислительные ресурсы. Цифровые двойники и большие данные (DICR-2019)», 2019. С. 139–147.

8. *Naumova V.V., Eremenko V.S., Platonov K.A., Dyakov S.V., Patuk M.I., Eremenko A.S.* Development of geographically distributed information-analytical geological environment // *Russian Journal of Earth Sciences*. 2019. Vol. 19, No. 6. DOI:10.2205/2019ES000696.

9. *Eremenko V.S., Naumova V.V., Platonov K.A., Dyakov S.E., Eremenko A.S.* The main components of a distributed computational and analytical environment for the scientific study of geological systems // *Russian Journal of Earth Sciences*. 2018. Vol. 18, no. 6 (current). DOI: 10.2205/2018ES000636

10. *Eremenko V.S., Naumova V.V.* Computational and Analytical Environment for Processing and Analysis of Geological Data // *Proceedings of the V International Conference "Information Technologies in Earth Sciences and Applications for Geology, Mining and Economy (ITES&MP-2019)"*, Moscow, Russia, October 14–18, 2019. Published on CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073). Vol-2527. Edited by Vera V. Naumova, Aleksandr S. Eremenko. P. 14–19.

11. *Tkaczyk D., Szostek P., Fedoryszak M., Dendek P., Bolikowski L.* CERMINE: automatic extraction of structured metadata from scientific literature. In *International Journal on Document Analysis and Recognition*. 2015. Vol. 18, no. 4. P. 317–335. DOI: 10.1007/s10032-015-0249-8.

12. *Maynard D., Bontcheva K., Augenstein I.* *Synthesis Lectures on the Semantic Web: Theory and Technology*, December 2016. Vol. 6, No. 2. P. 1–194.

DIGITAL REPOSITORY "GEOLOGYSOURCE.RU": OPEN ACCESS TO SCIENTIFIC PUBLICATIONS ON RUSSIAN GEOLOGY

Mikhail I. Patuk¹, Vera V. Naumova², Vitaliy S. Eremenko³

Vernadsky State Geological Museum RAS, Moscow (Russia)

¹patuk@mail.ru, ²naumova_new@mail.ru, ³vitaer@gmail.com

Abstract

The article describes new approaches related to the collection of data from heterogeneous information systems of access to scientific publications using open international standards and protocols for the formation of systems of open access to scientific geological publications. Based on developed and adapted approaches and technological solutions, a set of programs of information and analytical system of access to scientific publications has been implemented, implementing functions of collection, search, cataloguing, filtering and management of scientific publications and their metadata.

Keywords: *Information technology, Earth sciences, repository, scientific publications.*

REFERENCES

1. Budapest Open Access Initiative [electronic resource] URL: <https://www.budapestopenaccessinitiative.org> (date of the application: 23.05.2020).
2. Open Archive Initiative, OAI [electronic resource] URL: <https://openarchives.org/> (date of the application: 23.05.2020).
3. *Brase J.* DataCite – A global registration agency for research data // In: Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology. IEEE, Beijing. 2009. P. 257–261. DOI: 10.1109/COINFO.2009.66.
4. *Wilkinson M.D. et al.* The FAIR Guiding Principles for scientific data management and stewardship // In: *Sci. Data* 3:160018. 2016. P. 1. DOI: 10.1038/sdata.2016.18.
5. *Claudia Emerson, Elaine M. Faustman, Mustapha Mokrane, Sandy Harrison.* World Data System (WDS) Data Sharing Principles // Zenodo. 2015.

<http://doi.org/10.5281/zenodo.34354>.

6. Naumova V.V., Belousov A.V. Digital repository «Geology of the Russian Far East» – an open access to the spatially distributed online scientific publications // Russian Journal of Earth Sciences. 2014. Vol. 14, № 1. P. 1–8.

7. Naumova V.V., Platonov K.A., Eremenko V.S., Patuk M.I., Dyakov S.V. Informacionno-analiticheskaja sreda dlja podderzhki nauchnyh issledovanij v geologii: tekushhee sostojanie i perspektivy razvitija // Trudy XVII Mezhdunarodnoj konferencii «Raspredelennye informacionno-vychislitel'nye resursy. Cifrovye dvojniki i bol'shie dannye (DICR-2019)», 2019. S. 139–147.

8. Naumova V.V., Eremenko V.S., Platonov K.A., Dyakov S.V., Patuk M.I., Eremenko A.S. Development of geographically distributed information-analytical geological environment // Russian Journal of Earth Sciences. 2019. Vol. 19, No. 6. DOI:10.2205/2019ES000696.

9. Eremenko V.S., Naumova V.V., Platonov K.A., Dyakov S.E., Eremenko A.S. The main components of a distributed computational and analytical environment for the scientific study of geological systems // Russian Journal of Earth Sciences. 2018. Vol. 18, no. 6 (current). DOI: 10.2205/2018ES000636.

10. Eremenko V.S., Naumova V.V. Computational and Analytical Environment for Processing and Analysis of Geological Data // Proceedings of the V International Conference “Information Technologies in Earth Sciences and Applications for Geology, Mining and Economy (ITES&MP-2019)”, Moscow, Russia, October 14–18, 2019. Published on CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073). Vol-2527. Edited by Vera V. Naumova, Aleksandr S. Eremenko. P. 14–19.

11. Tkaczyk D., Szostek P., Fedoryszak M., Dendek P., Bolikowski L. CERMINE: automatic extraction of structured metadata from scientific literature. In International Journal on Document Analysis and Recognition. 2015. Vol. 18, no. 4. P. 317–335. DOI: 10.1007/s10032-015-0249-8.

12. Maynard D., Bontcheva K., Augenstein I. Synthesis Lectures on the Semantic Web: Theory and Technology, December 2016. Vol. 6, No. 2. P. 1–194.

СВЕДЕНИЯ ОБ АВТОРАХ



ПАТУК Михаил Иванович – к. г.-м. н., и. о. н. с., Научный отдел Государственного геологического музея им. В.И. Вернадского РАН, Москва, Россия.

Michail I. PATUK – PhD, scientific department, Vernadsky State Geological Museum RAS, Moscow (Russia).

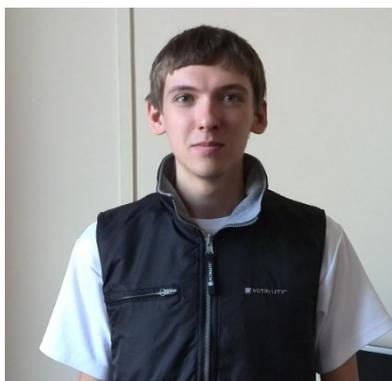
Email: patuk@mail.ru



НАУМОВА Вера Викторовна – д. г.-м. н., г. н. с., зав. Научным отделом Государственного геологического музея им. В.И. Вернадского РАН, Москва, Россия.

Vera V. NAUMOVA – Prof., head of SGM scientific department, Vernadsky State Geological Museum RAS, Moscow (Russia).

Email: Naumova_new@mail.ru



ЕРЁМЕНКО Виталий Сергеевич – и. о. м. н. с., Научный отдел Государственного геологического музея им. В.И. Вернадского РАН, Москва, Россия.

Vitaliy S. EREMENKO – scientific department, Vernadsky State Geological Museum RAS, Moscow (Russia).

Email: vitaer@gmail.com

Материал поступил в редакцию 28 мая 2020 года