

Информационные технологии в системном анализе

Подход к хранению и обработке первичной слабоструктурированной геолого-геофизической информации*

А.Л. ЦАРЕГОРОДЦЕВ^{1,II}, С.С. СЛОБОДЕНЮК^I, С.В. ВОЛОШИН^I, А.В. ВОХМИНЦЕВ^{I,III}

^I Автономное учреждение «Югорский научно-исследовательский институт информационных технологий», г. Ханты-Мансийск, Россия

^{II} Югорский государственный университет, г. Ханты-Мансийск, Россия

^{III} Челябинский государственный университет, г. Челябинск, Россия

Аннотация. Методы хранения производственных данных являются важным фактором информатизации нефтяных месторождений. При этом возникают некоторые проблемы их использования, из-за несовместимости форматов, дублирования информации, скорости манипулирования данными и т.д. В данной статье описан начальный этап разработки автоматизированной информационной системы, являющейся единым информационным пространством результатов исследований первичной геолого-геофизической информации. Представлены методы структурирования и предварительной обработки разнородной геолого-геофизической информации, относящейся к описанию скважин, их местоположению, эксплуатационным показателям, результатам инклинометрии и каротажным данным, а также к пересечению конкретных геологических пластов, с целью ее дальнейшего хранения в реляционной базе данных и обработки методами статистического анализа и машинного обучения для прогнозирования перспективных интервалов разреза и выдачи рекомендаций по нефтедобыче на основе комплексного анализа. Также представлена и реализована реляционная модель для структурированного хранения информации, обеспечивающая высокую скорость операций манипулирования данными.

Ключевые слова: *разнородная геолого-геофизическая информация, неструктурированные данные, реляционная модель, первичная обработка данных.*

DOI: 10.14357/20790279190401

Введение

Актуальность организации хранения и автоматической обработки больших объемов разнородной геолого-геофизической информации, в частности, результатов геофизических исследований скважин и результатов их испытаний, в реляционной базе данных (далее БД), полученных с нефтяных месторождений, обусловлена тем, что нефтедобывающие

предприятия повторно используют их для планирования своей дальнейшей работы. При этом структурирование разнородных геолого-геофизических и промысловых данных является весьма трудозатратным процессом, т.к. геолого-геофизические данные имеют большое количество разнородных параметров. Поэтому систематизированное хранение геолого-геофизической и промысловой информации, с целью ее дальнейшей обработки, анализа закономерностей, построения моделей и прогнозирования

* Данная работа была выполнена при поддержке Российского фонда фундаментальных исследований проект № 19-47-890001

результатов эксплуатации является актуальной задачей. Решение указанной задачи является первоначальным этапом обширной работы по разработке автоматизированной информационной системы «Цифровая комната данных» для формирования упорядоченной геолого-геофизической информации, направленной на повышение эффективности разработки нефтяных месторождений баженовско-абалакской свиты на основе методов глубокого машинного обучения.

Задачи обработки разнородных данных возникли ранее и возникают до сих пор в самых разных областях. Например, в здравоохранении [1], экологии [2] и экономической безопасности [3]. При этом работ, связанных с обработкой и хранением геолого-геофизических данных, не так много. В частности, в настоящее время было предложено несколько информационных систем (далее ИС) и централизованных баз данных для сбора и анализа геофизических данных [4-8]. В работе [4] описана горно-геологическая информационная система для хранения и обработки всего объема данных о рудовмещающем горизонте, применяемых для метода скважинного подземного выщелачивания. Эта ИС состоит из пяти модулей, каждый из которых решает отдельные задачи. Достоинствами описанной системы является автоматическая обработка результатов геофизического исследования скважин в стандарте LAS-файлов, формирование единой реляционной и непротиворечивой базы данных, а также наличие клиентского программного обеспечения для визуализации и редактирования имеющихся данных. Из существенных недостатков можно отметить отсутствие возможности автоматической обработки файлов, отличных от LAS.

В работах [6] и [7] описывается концептуальная модель и разработанное на ее основе программное обеспечение для хранения комплексных геофизических данных. Особенность концепции заключается в обеспечении оперативного доступа к геолого-геофизическим материалам и приведение их к единому формату для удобного хранения и использования. Единственным достоинством описанной модели можно выделить обеспечение структурированного и централизованного хранения первичной сейсморазведочной информации. В разработанном продукте отсутствует обработка хранящейся информации, что затрудняет проведение ее анализа.

Настоящая работа описывает методы структурирования обобщенной геолого-геофизической информации, полученной при поисково-разведочных и эксплуатационных работах на месторождениях Ханты-Мансийского автономного округа. Полученная в результате применения описанных

методов реляционная модель данных позволит эффективно разрабатывать и применять алгоритмы и методы статистического анализа и машинного обучения для повышения эффективности разработки месторождений баженовско-абалакского комплекса. Для применения методов машинного обучения к данным необходима их тщательная обработка: устранение противоречий, дубликатов, отклонений от ожидаемого распределения и т.д. Разработанное хранилище позволяет использовать данные напрямую, от исследователя требуется лишь сформировать сбалансированную выборку, сделать нормализацию необходимых признаков.

Например, полученная в рамках работы модель данных позволяет (при наличии необходимого объема данных) прогнозировать перспективные (с точки зрения дебита нефти) интервалы разреза на основе методов машинного обучения по данным каротажных исследований и результатов испытаний скважин. В дальнейшем, с использованием описанных методов, планируется разработка системы выдачи рекомендаций по нефтедобыче на основе комплексного анализа геолого-геофизической информации.

1. Данные и методы проверки их корректности

Большая часть всей информации, связанная со скважинами (месторождение, пластопересечения, категория скважины и т.д.) и ее эксплуатационными характеристиками, предоставляется в произвольном и неструктурированном виде, таком как электронные таблицы или текстовое описание, что в свою очередь затрудняет процесс автоматической обработки подобных данных. Помимо этого, имеются большие объемы данных, полученных из результатов проведения геофизических исследований скважин и представленных LAS-файлами [9]. Также, в специализированных стандартах inc и dev, имеются результаты проведения исследований траектории стволов скважин (инклинометрия).

На данном этапе, для автоматизации обработки больших объемов данных и приведении их к необходимой структуре, было разработано программное обеспечение, позволяющее преобразовывать результаты геофизического исследования скважин в стандартах las, inc, dev, csv файлов к соответствующей структуре модели, с его последующей загрузкой в БД. Например, на рисунке 1 показана блок-схема алгоритма обработки inc файлов и сохранения приведенной к требуемой структуре БД информации о проведении мероприятий по инклинометрии скважин и полученные при этом результаты исследования с установленными необ-

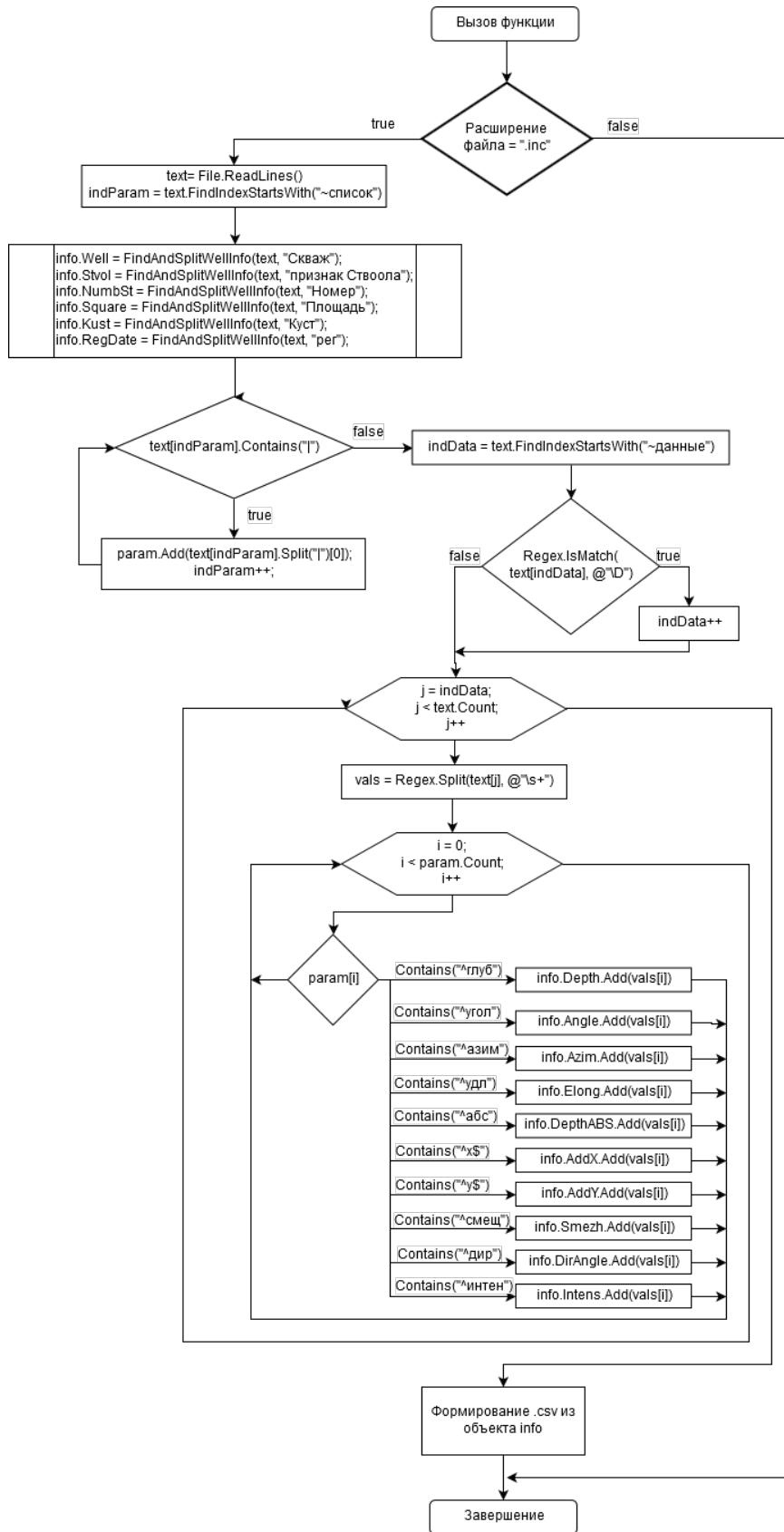


Рис. 1. Блок-схема алгоритма обработки файлов с результатами инклинометрии

ходимыми связями. Для заполнения временными данными соответствующего объекта применяется поиск текста с помощью регулярных выражений.

В настоящее время не реализована автоматическая обработка эксплуатационных показателей скважин и информации о пластопересечениях из-за неструктурированного источника данных. Поэтому необходима их ручная предварительная обработка и приведение к определенному формату для загрузки в базу данных.

Для возможности дальнейшего использования данных и получения более точных результатов при их обработке методами глубокого машинного обучения, необходимо решить ряд задач, связанных с проверкой корректности загружаемых данных.

Первой задачей, которая возникает из-за несоответствия единиц измерения некоторых значений характеристики, является приведение всех значений определенного показателя к единой системе измерения. К примеру, для этого применялись дополнительные расчеты по формуле (1) для преобразования добываемых объемов нефти из $m^3/сут$ в $t/сут$, с учетом средней плотности нефти для конкретного месторождения.

$$m = v * \rho,$$

Кроме того, наиболее трудозатратным было выполнение преобразования всех используемых координат к единой системе координат Pulkovo 1942. Для этого вокруг каждого месторождения была определена площадь ограничивающая расположение скважин. Если координаты не попадают ни в одну из указанных областей, то последовательно подбираем исходную систему координат.

Например, для преобразования значения долготы из координатной системы WGS 1984 в Pulkovo 1942, согласно ГОСТ 51794-2008 [10], используются расчеты по формуле Бурса-Вольфа:

$$Lon = \frac{ro}{\left(\left(a * \left(1 - e2 * \sin\left(Lat * \frac{\pi}{180} \right)^2 \right)^{-0.5} + H \right) * \cos\left(Lat * \frac{\pi}{180} \right) \right) * \left(-\Delta X * \sin\left(Lon * \frac{\pi}{180} \right) + \Delta Y * \cos\left(Lon * \frac{\pi}{180} \right) \right)}, \quad (2)$$

где $ro, a, e2, \Delta X, \Delta Y$ – константные и вычисляемые параметры, Lat, Lon, H – входные параметры широты, долготы и высоты соответственно. Блок-схема полного алгоритма преобразования координат показана на рисунке 2. Для данного преобразования используются три параметра трансформации (dx, dy, dz), остальные параметры равны 0. При этом, все угловые значения передаются и возвращаются в десятичных градусах, а высоты – в метрах. После чего проверяем

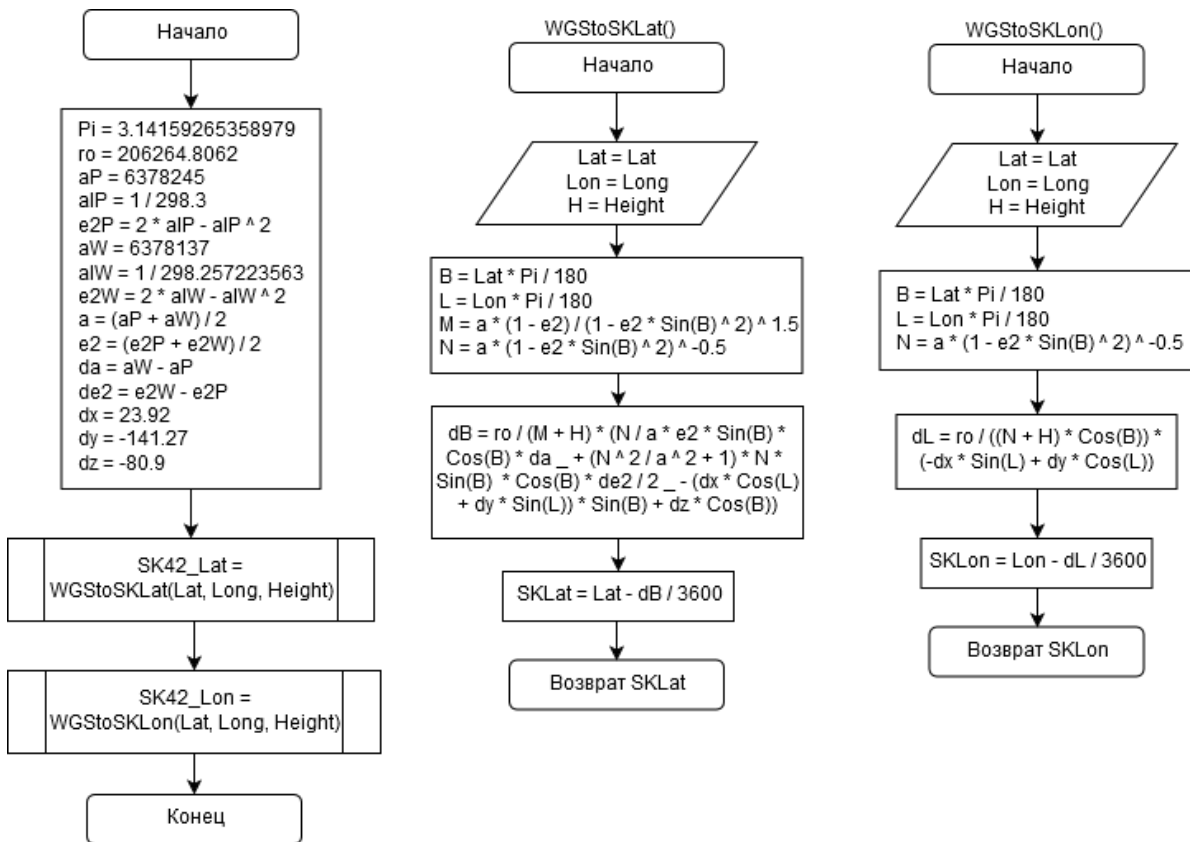


Рис. 2. Алгоритм преобразования координат из системы WGS 1984 в Pulkovo 1942.

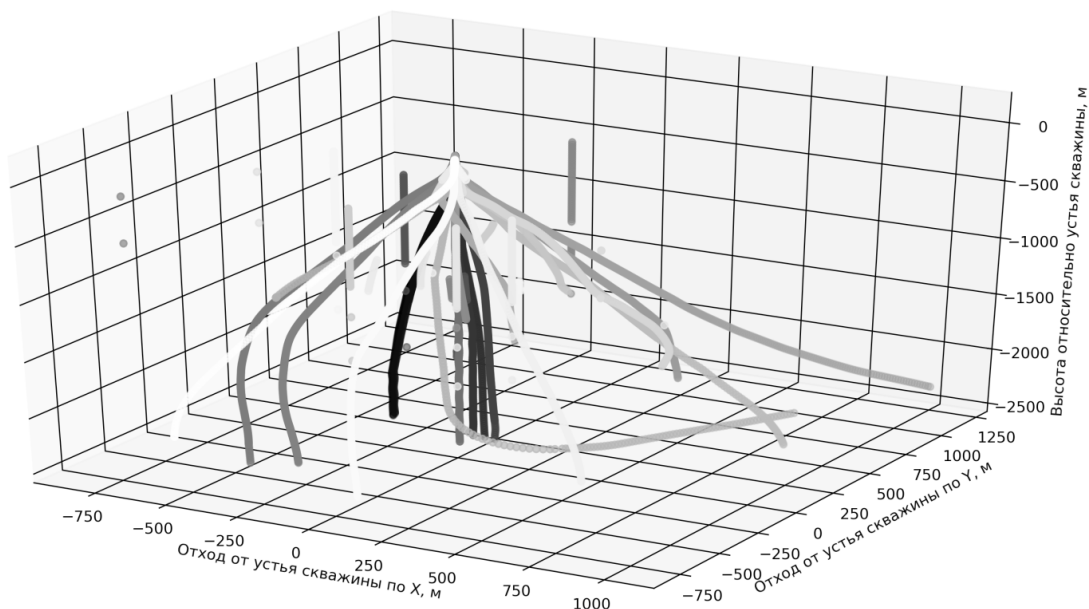


Рис. 3. Визуализация исходных данных инклинометрии

полученные значения координат на принадлежность размеченным областям.

Следующей проблемой являлось наличие в результатах инклинометрии противоречивой информации, описывающей траектории и координатные привязки стволов скважин. Для решения данной задачи необходимо было провести сопоставление, анализ и поиск отклонений таких показателей, как координаты устья и забоя стволов, несоответствие указанной длины ствола с рассчитанной по координатам из результата инклинометрии, а также проведение поиска стволов скважин,

у которых абсолютная глубина забоя превышает фактическую длину ствола.

Еще одной задачей при проверке корректности результатов инклинометрии было выявление разрывов в траекториях стволов скважин, а также поиск траекторий, имеющих координаты начала отличные от координат устьев. Для этого была построена и проанализирована визуальная модель траекторий стволов. На рисунке 3 и 4 представлена визуализация исходных и обработанных данных по некоторым скважинам соответственно. Различия между данными для визуализации рисунков 3 и 4 заключаются в устране-

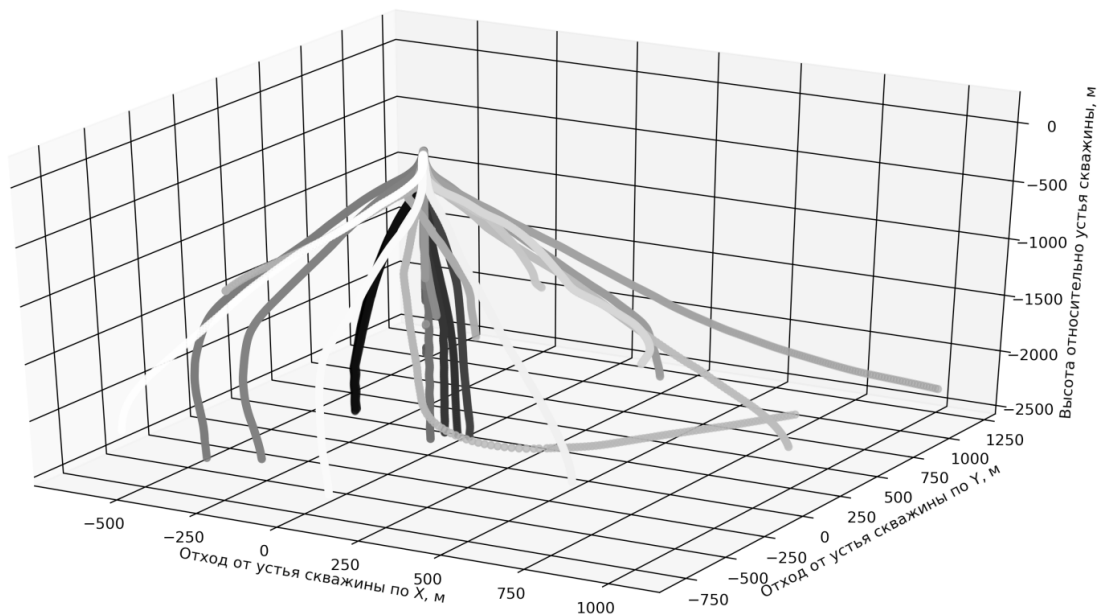


Рис. 4. Визуализация обработанных данных инклинометрии

нии разрывов: расстояние между соседними точками сравнивалось с шагом исследования. Ствол скважины восстанавливался в случаях, когда аномальные изменения происходили в одной из координат. Так как в иных случаях однозначно восстановить ствол скважины невозможно, ошибка не устранялась, однако такие скважины получали отдельную отметку о содержащейся ошибке, начиная с некоторой глубины.

Решение указанных задач позволило предварительно проверить имеющиеся данные, исправить найденные ошибки и загрузить в базу более корректные значения, обеспечивающие наиболее качественную выборку данных для последующей обработки методами машинного обучения.

2. Модель базы данных

Как известно, для представления данных в реляционной базе данных используются понятия множества сущностей и связь. При этом реляционная модель применяет только один способ организации данных через отношения, которые реализуют множество сущностей и связей между ними. Этот подход с одной стороны позволяет

не учитывать особенности данных, хранящихся в реляционной базе данных, с другой предполагает точное знание структуры хранимых данных. При использовании реляционной модели хранения данных все строки в какой-либо таблице всегда имеют одинаковый набор полей, кроме того, все названия столбцов, первичные и внешние ключи в таблицах должны быть определены заранее. Наш подход к хранению и обработке первичной слабоструктурированной геолого-геофизической информации не подразумевает изменение структуры хранения самих данных и их типов. Кроме того, массив геолого-геофизических данных, который необходимо обрабатывать планируется большим (1-2 Тб).

С другой стороны большое количество транзакции на доступ и изменение геолого-геофизических данных, требует соответствия базы данных требованиям ACID (Atomicity, Consistency, Isolation, Durability), что возможно, только в реляционных базах данных.

Оба эти критерия стали решающим фактором при определении подхода к хранению и обработке первичной слабоструктурированной геолого-геофизической информации.

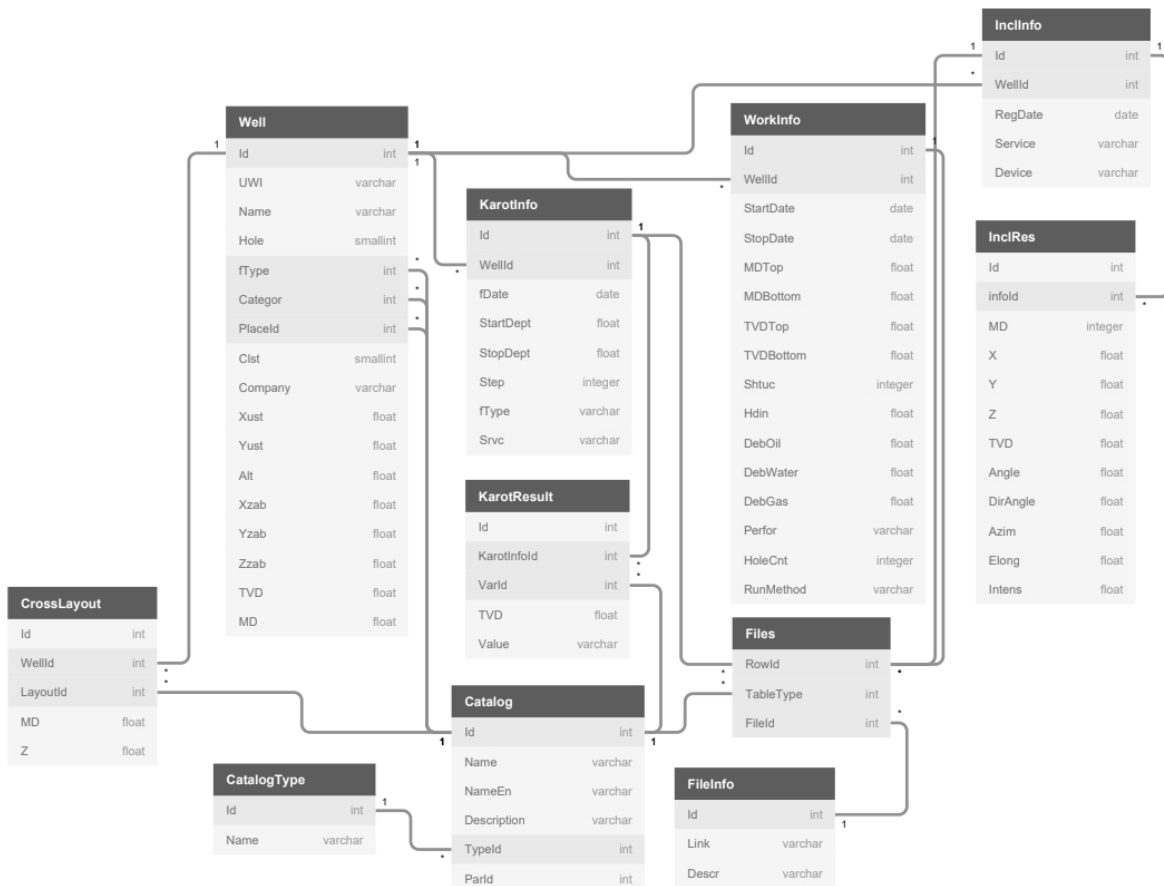


Рис. 5. Структура БД

После анализа всех имеющихся видов информации было спроектировано универсальное структурированное хранилище геолого-геофизических данных, приведенное к третьей нормальной форме и обеспечена целостность и согласованность хранимых данных [11, 12]. Структура разработанной реляционной базы данных представлена на рисунке 5.

Центральным звеном спроектированной схемы является таблица *Well*, которая содержит информацию о стволах скважин, включая название скважины и ее тип, месторождение, куст, компанию – владельца, альтитуду, координаты устья и забоя. Отдельно для каждого ствола хранится информация о его пластопересечениях с указанием абсолютного значения глубины. Также для каждого ствола скважины может храниться несколько наборов данных, связанных с интервалами и результатами испытания (табл. *WorkInfo*), проведением инклинометрических и каротажных исследований (табл. *InclInfo*, *IncRes* и *KarotInfo*, *KarotResult* соответственно). В силу того, что результаты каротажных исследований являются квазиструктурированными данными, поскольку они имеют неопределенный набор измеряемых параметров, которые могут меняться с течением времени, было принято решение хранить их в отдельной таблице *Catalog*. Организация структуры такого типа для хранения результатов каротажных исследований является наиболее оптимальным вариантом для хранения квазиструктурированных данных [13]. Помимо этого, имеется возможность указать мнемоники (синонимы) данных параметров, ссылаясь на родителя в поле *ParId*. При необходимости, имеется возможность обратиться к файлу – первоисточнику, из которого была извлечена информация. В БД хранятся ссылки на расположенные в каталогах системы исходные файлы и устанавливаются связи между ними и записями таблиц, содержащих извлеченную из данных файлов информацию.

Совокупность хранения подобной информации в таком виде дает возможность для дальнейшей разработки программного продукта, обеспечивающего оперативный доступ ко всей необходимой информации, ее обработку, визуализацию, анализ, прогнозирование и т.д.

Поскольку базу данных планируется использовать для хранения и обработки большого массива данных (~ 1-2Тб), при этом сами SQL-запросы к данным не всегда являются тривиальными, пришлось оптимизировать SQL-запросы. Для выявления проблемных областей использовали план

выполнения запросов. Были добавлены простые индексы для часто используемых полей, таких как координаты и длина ствола скважины, параметры каротажных исследований, дебит нефти и т.д., а также созданы индексы по выражениям, обеспечивающие более быструю обработку вычисляемых значений в запросе.

Таким образом, разработанная модель реляционной базы данных для хранения геолого-геофизической информации и методы ее структурирования являются первой стадией разработки автоматизированной информационной системы «Цифровая комната данных», являющаяся единым информационным пространством результатов исследований первичной геолого-геофизической информации в сфере освоения трудноизвлекаемых запасов. Целью которого является рентабельная разработка баженовско-абалакской свиты за счет использования отечественного комплекса технологий и повышения эффективности разработки за счет снижения Удельного показателя – отношения затрат на строительство скважины к ее накопленной добыче нефти.

Заключение

В результате проделанной работы была спроектирована реляционная база данных, предназначенная для хранения и обработки разнородной геолого-геофизической информации. Разработано программное обеспечение, позволяющее преобразовывать результаты геофизического исследования скважин в стандартах *las*, *inc*, *dev*, *csv* файлов к соответствующей структуре модели данных. Проведено сопоставление, анализ и поиск противоречивой информации описывающей траектории, и координатных привязок стволов скважин. Решение указанных задач позволило предварительно проверить имеющиеся данные, исправить найденные ошибки и загрузить в базу более корректные значения, обеспечивающие наиболее качественную выборку данных для последующей обработки методами машинного обучения.

В дальнейшем планируется провести работу по оптимизации алгоритмов обработки слабоструктурированных *las*, *inc* и *dev* файлов, реализовать дополнительные методы проверки корректности промысловых и полученных из результатов каротажных исследований данных, минимизировать взаимодействие с пользователем при обработке неструктурированных данных, разработать клиентское программное обеспечение и расширить схему базы данных для хранения данных сейсморазведки.

Литература

1. Царегородцев А.Л. Особенности хранения и передачи частично структурированной медицинской информации в АИС // Вестник Новосибирского гос. ун-та. Серия: Информационные технологии. – 2010. – Т. 8. № 2, с.65-69.
2. Кочергин Г.А., Зыков А.С., Царегородцев А.Л., Горлов Н.В. Автоматизированная информационная система мониторинга изменений на основе космических снимков // Известия Алтайского государственного университета. 2017. № 4 (96). С. 78-83.
3. Митяков Е.С., Митяков С.Н. Алгоритм импорта и первичной обработки данных при анализе экономической безопасности регионов России // Современные проблемы науки и образования. – 2015. – № 1-1.
4. Истомин А.Д., Носков М.Д., Чеглоков А.А. Горно-геологическая информационная система для месторождения, разрабатываемого методом скважинного подземного выщелачивания // ГИАБ. 2012. №7, с.178-183.
5. Кобен Г.К. Формирование единой базы данных геолого-геофизической информации по нефтегазовым месторождениям Казахстана / Г.К. Кобен, А.М. Турсынова ; науч. рук. Г.А. Жылкыбаева // Проблемы геологии и освоения недр : труды XXI Международного симпозиума имени академика М.А. Усова студентов и молодых ученых, посвященного 130-летию со дня рождения профессора М. И. Кучина, Томск, 3-7 апреля 2017 г. : в 2 т. – Томск : Изд-во ТПУ, 2017. – Т. 1. – [С. 372-374].
6. Королева А.В. Программное обеспечение для организации хранения массовых геофизических данных // Вестник ИрГТУ. 2010. №2 (42), с.11-15.
7. Пешикова А.А., Мироманов А.В. Концептуальная модель геоинформационного проекта для хранения комплексных геофизических данных и управления ими // Известия СО РАН. Геология, поиски и разведка рудных месторождений. 2017. №4 (61), с.77-85.
8. Shaowei Huang, Chao Yang, Dandan Wang. A Storage Platform of Petroleum Production based on MongoDB // International Conference on Computer Science and Application Engineering, 2017, с 496-503.
9. Стандарт хранения данных каротажа LAS (версии 1.2 и 2.0) Учебное пособие / Д.К.Нурғалиев, В.Е. Косарев, Л.Р. Косарева, Б.М. Насыртдинов – Казань: Казанский университет, 2012. – 23 с.
10. ГОСТ Р 51794-2008. Глобальные навигационные спутниковые системы. Системы координат. Методы преобразований координат определяемых точек. – Москва: Стандартинформ, 2009. – 25 с.
11. Корнеева Е.В. Методика проектирования данных на основе слабоструктурированной информации // Сервис в России и за рубежом. 2009. №2.
12. Костикова А.В. Проектирование баз данных на основе слабоструктурированной информации // Научное сообщество студентов XXI столетия. ТЕХНИЧЕСКИЕ НАУКИ: сб. ст. по мат. LXIV междунар. студ. науч.-практ. конф. № 4(63).
13. Палей Д. Моделирование квазиструктурированных данных. //Открытые системы, 2002, № 9 С. 57.

Царегородцев Александр Леонидович. Автономное учреждение «Югорский научно-исследовательский институт информационных технологий» г. Ханты-Мансийск, Россия. Первый заместитель директора, кандидат технических наук, доцент. Количество печатных работ: 20. Область научных интересов: базы данных, извлечение знаний, системный анализ, электронные социальные сети. Email: alexts01@yandex.ru

Слободенюк Сергей Сергеевич. Автономное учреждение «Югорский научно-исследовательский институт информационных технологий» г. Ханты-Мансийск, Россия. Программист. Количество печатных работ: 1. Область научных интересов: математическая статистика, программная реализация интеллектуальных систем, нейронные сети, обработка естественного языка. Email: slobodenykss@uriit.ru

Волошин Семен Викторович. Автономное учреждение «Югорский научно-исследовательский институт информационных технологий» г. Ханты-Мансийск, Россия. Ведущий программист. Количество печатных работ: 2. Область научных интересов: математическая статистика, программная реализация интеллектуальных систем, нейронные сети, обработка естественного языка. Email: VoloshinSV@uriit.ru

Вохминцев Александр Владиславович. ФГБОУ ВО «Челябинский государственный университет». г. Челябинск, Россия. Заведующий научно-исследовательской лабораторией «Интеллектуальные информационные технологии и системы». Количество печатных работ: 39. Область научных интересов: цифровая обработка информации, методы сопоставления и регистрации изображений, методы одновременной навигации и составления карты в неизвестном пространстве, распознавание личности по лицу. E-mail: vav@csu.ru

An approach to the storage and processing of original semi-structured geological and geophysical information

A.L. Tsaregorodtsev¹, S.S. Slobodenyuk¹, S.V. Voloshin¹, A.V. Vokhmintcev^{1,II}

¹ Ugra Research Institute of Information Technologies

^{II} Chelyabinsk State University

Abstract. Methods of storage of production data are an important factor in the Informatization of oil fields. In this case, there are some problems of their use, due to the incompatibility of formats, duplication of information, speed of data manipulation, etc. This article describes the initial stage of development of an automated information system, which is a single information space of the results of research of primary geological and geophysical information. The methods of structuring and pre-processing of heterogeneous geological and geophysical information related to the description of wells, their location, performance indicators, the results of inclinometry and logging data, as well as to the intersection of specific geological formations, for its further storage in a relational database and processing methods of statistical analysis and machine learning to predict long-term intervals of the section and issue recommendations for oil production based on a comprehensive analysis. A relational model for structured data storage, providing high speed of data manipulation operations, is also presented and implemented.

Keywords: *heterogeneous geological and geophysical information, unstructured data, relational model, primary data processing.*

DOI: 10.14357/20790279190401

References

1. *Tsaregorodtsev, A.L.* 2010. Osobennosti hraneniya i peredachi chastichno strukturirovannoi meditsinskoy informatsii v AIS [Features of storage and transfer of partially structured medical information in information system]. Vestnik Novosibirskogo gosudarstvennogo universiteta [Bulletin of Novosibirsk State University]. 8(2):65-69
2. *Kochergin, G.A., Zykov, A.S., Tsaregorodtsev, A.L., Gorlov, N.V.* 2017. Avtomatizirovannaya informatsionnaya sistema monitoring izmeneniy na osnove kosmicheskikh snimkov [Automated information system for monitoring changes based on satellite images]. Izvestiya Altaiskogo gosudarstvennogo universiteta [Altaic State University News]. 4(96):78-83.
3. *Mityakov, E.S., Mityakov, S.N.* 2015. Algoritm importa i pervichnoy obrabotki danih pri analize ekonomicheskoy bezopasnosti regionov Rossii [Algorithm of import and original data processing in the analysis of economic security of Russian regions]. Sovremennye problemi nauki i obrazovaniya [Modern problems of science and education]. 1-1 Available at: <https://www.science-education.ru/ru/article/view?id=17764> (accessed September 16, 2019).
4. *Istomin, A.D., Noskov, M.D., Cheglov, A.A.* 2012. Gorno-geologicheskaya informatsionnaya sistema dlya mestorojdeniya, razrabativaemogo metodom skvajinnogo podzemnogo vishelachivaniya [Mining and geological information system for a field developed by borehole underground leaching]. GIAB. 7:178-183.
5. *Koben, G.K., Tursinova, A.M., Jilkivbaeva, G.A.* 2017. Formirovanie edinoi bazi danih geologo-geofizicheskoy informatsii po neftegazovim mestorojdeniyam Kazahstana [Formation of a unified database of geological and geophysical information on oil and gas fields of Kazakhstan]. Problemi geologii i osvoeniya nedr: trudi XXI Mejdunarodnogo simpoziuma imeni akademika M.A. Usova studentov I molodih uchenih, posvzhen-nogo 130-letiyu so dnya rojdeniya professor M.I. Kuchina [Problems of Geology and Subsoil Development: Proceedings of the XXI International Academician M. Usov Symposium for Students and Young Scientists, Dedicated to the 130th Birthday of Professor M. I. Kuchin]. 1:372-374.
6. *Korolyeva, A.V.* 2010. Programmnoe obespechenie dlya organizatsii hraneniya massovih geofizicheskikh danih [Software for organizing the storage of mass geophysical data]. Vestnik IrGTU [Bulletin of Irkutsk State Technical University]. 2(42):11-15.
7. *Peshkova, A.A., Miromanov, A.V.* 2017. Kontseptual'naya model' geoinformatsionnogo proekta dlya hraneniya kompleksnih geofizicheskikh danih i upravleniya imi [A conceptual model of a geographic information project for storing and managing complex geophysical data]. Izvestiya SO RAEN. Geologiya, poiski i razvedka rudnih mestorojdeniy [zvestiya SB RANS. Geology, prospecting and exploration of ore deposits]. 4(61):77-85.
8. *Shaowei, H., C. Yang, D. Wang.* 2017. A Storage Platform of Petroleum Production based on MongoDB. International Conference on Computer Science and Application Engineering. 496-503.

9. *Nurgaliev, D.K., Kosarev, V.E., Kosareva, L.R., Nasritdinov, B.M.* 2012. Standart hraneniya danih karotaja LAS (versii 1.2 i 2.0) Uchebnoje posobie [LAS Log Data Storage Standard (versions 1.2 and 2.0) Tutorial]. Kazan: Kazan University. 23p.
10. GOST R 51794-2008. 2009. Global'niye navigatsionnie sputnikovye sistemi. Sistemi koordinat. Metodi preobrazovaniy koordinat opredelyaemih toчек [Global navigation satellite systems. Coordinate systems. Methods of transforming coordinates of defined points]. Moscow: StandardinformPubls. 25 p.
11. *Korneeva, E.V.* 2009. Metodika proektirovaniya danih na osnove slabostruktirovannoi informatsii [Methodology for designing data based on semi-structured information]. Servis v Rossii i za rubezhom [Service in Russia and abroad]. 2.
12. *Kostikova, A.V.* 2018. Proektirovanie baz danih na osnove slabostruktirovannoi informatsii [Database design based on semi-structured information]. Nauchnoye soobshestvo studentov XXI stoletiya. TEHNICHESKIE NAUKI: sb. st. po mat. LXIV mejdunar. stud. nauch.-prakt. konf. [The scientific community of students of the XXI century. TECHNICAL SCIENCES: Sat. Art. by mat. LXIV Int. Stud. scientific-practical conf.]. Novosibirsk. 4(63).
13. *Paley, D.* 2002. Modelirovanie kvazistruktirovannih danih. [Modeling quasi-structured data.]. Otkritie sistemi [Open systems]. 9:57.

Tsaregorodtsev A.L. PhD, Ugra Research Institute of Information Technologies, 151 Mira str., Khanty-Mansiysk, 628011, Russia, e-mail alexts01@yandex.ru.

Slobodenyuk S.S. Ugra Research Institute of Information Technologies, 151 Mira str., Khanty-Mansiysk, 628011, Russia, e-mail slobodenykss@uriit.ru.

Voloshin S.V. Ugra Research Institute of Information Technologies, 151 Mira str., Khanty-Mansiysk, 628011, Russia, e-mail VoloshinSV@uriit.ru

Vokhmintcev A.V. Chelyabinsk State University, 129 Bratiev Kashirinykh st., 454001 Chelyabinsk, Russia, e-mail vav@csu.ru