

УДК 550.8.028

**СОЗДАНИЕ УНИВЕРСАЛЬНОГО
ХРАНИЛИЩА ДАННЫХ ПРОМЫСЛОВО
ГЕОФИЗИЧЕСКИХ ИССЛЕДОВАНИЙ**

А.В. Марков, Д.Е. Черкунов*, В.С. Шерстнёв

Томский политехнический университет
*ОАО «ТомскНИПИнефть», г. Томск
E-Mail: markovav@tpu.ru**Марков Александр Валерьевич**, магистрант кафедры вычислительной техники Института кибернетики ТПУ.

E-mail: markovav@tpu.ru

Область научных интересов: хранилища данных, разработка ПО.

Черкунов Дмитрий Евгеньевич, заведующий сектором технологий баз данных ОАО «ТомскНИПИнефть», г. Томск.

E-mail: CherkunovDE@

nipineft.tomsk.ru

Область научных интересов: хранилища данных, разработка ПО.

Шерстнёв Владислав Станиславович, доцент кафедры вычислительной техники Института кибернетики ТПУ.

E-mail: vss@tpu.ru

Область научных интересов: разработка ПО.

Показана возможность применения информационно-технологических подходов в области систематизации и обработки данных промыслово-геофизических исследований с целью анализа выработки запасов и автоматизации процесса анализа выработки запасов. Описана общая концепция создания хранилища данных для работы с данными промыслово-геофизических исследований.

Ключевые слова:

Промыслово-геофизические исследования, хранилища данных, геолого-технические мероприятия.

Key words:

Downhole logging, Data Warehouse, geological and technical events.

В настоящее время, в связи с интенсивным развитием отечественного нефтегазового комплекса, все большую актуальность у компаний-недропользователей приобретает проблема освоения залежей с трудно извлекаемыми запасами углеводородов. Эффективность освоения трудно извлекаемых запасов в районах со сложной геологией выдвигает все более жесткие требования к точности и достоверности построения геологических моделей.

Эффективность освоения залежей в районах со сложной геологией выдвигает все более жесткие требования к точности и достоверности построения геологических моделей.

Ключевым фактором, влияющим на достоверность модели, является степень изученности геологического строения и параметров среды, что в свою очередь зависит от наличия и полноты использования геолого-геофизических, геолого-промысловых, сейсморазведочных данных, материалов исследований керна.

Достоверность геологической модели и качественный сбор исходных данных оказывает значительное влияние на качество работ по проектированию и мониторингу разработки месторождений, от эффективности и качества выполнения которых, в свою очередь, зависят сроки освоения месторождения и начало добычи углеводородного сырья (нефти и газа) [1].

Показатели эффективности и качества разработки, в свою очередь, напрямую зависят от полноты и качества исходных материалов, то есть всего комплекса проводимых на месторождении исследований.

Одними из основных видов исследований является промыслово-геофизические исследования (ПГИ). Геологи и разработчики месторождений с помощью ПГИ решают следующие основные задачи:

- мониторинг и контроль за разработкой месторождений;
- построение геолого-технологической модели месторождения;

- планирование геолого-технологических мероприятий на скважинах;
- расчет степени участия нефтенасыщенных пропластков в общем дебите скважины [2].

Весь комплекс указанных мероприятий основан на использовании актуальных, систематически пополняемых данных единой информационной системы.

Наряду с решением задачи информационного обеспечения конечных пользователей (доступа к актуальным данным ПГИ) посредством использования данной системы, обеспечивается взаимодействие различных подразделений в процессе сбора, подготовки, верификации и загрузки данных ПГИ.

Схематично этот процесс приведен на рис. 1. Подрядчик на промысле выполняет исследования и выдает результаты в формализованной единой форме недропользователю. Недропользователь загружает данные в систему по средствам веб-инструментов загрузки. Единая структура входных данных и разработанные программы-загрузчики обеспечивают автоматизацию данного процесса. Ответственные специалисты по промыслово-геофизическим исследованиям со стороны недропользователя через интерфейс системы выполняют верификацию всех вновь загруженных исследований. Согласованные результаты исследования становятся доступны геологическим службам недропользователя. По несогласованным результатам исследования строится отчет и направляется на доработку подрядчику.



Рис 1. Жизненный цикл ПГИ

На первом этапе подрядчик должен заполнить отчеты в строго регламентированной форме в соответствии с регламентом, в котором оговариваются структура файлов отчетов, форматы и порядок именования. Для облегчения работы подрядчиков, разработаны типизированные шаблоны отчетов в формате Microsoft Office 2010 (xlsx).

Для дальнейшего автоматического разбора и систематизации поступивших отчетов и загрузки данных в хранилище данных (ХД) необходимо точное заполнение шаблонов в соответствии с регламентом. Для этого в шаблоны внедрены механизмы выбора дат, списки доступных значений для определенных параметров исследований, а также защита от ввода некорректных данных в произвольные области отчетов.

Пример отчета приведен на рис. 2.

D29		f									
	A	B	C	D	E	F	G	H	I	J	
1	Заказчик	ОАО "Томскнефть" ВНК									
2	Исполнитель	ООО "Томскнефтегазгеофизика"									
3	Реквизиты	Тел. (38259) вед. геоф. КИЭ, 6-34-37 нач. КИЭ. E-mail: geofizi@strj.tomsk.su									
4											
5	Интерпретатор	ведущий геофизик В.А. Яхина									
6	Начальник партии	Свиридов В.									
7											
8	Отчет по результатам обработки данных ШГИ										
9	Скважина №	191									
10	Куст №	3									
11	Месторождение	Крапивинское									
12									Согласовано	Комментарий	
13	Цель исследования	определение				профиль приемистости		да			
14						техническое состояние эк		да			
15						источник обводнения		да			
16						контроль перфорации		да			
17											
18	Пласт	Ю1(2+3а+3б)									
19	Дата интерпретации	16	1	2011							
20	Дата исследования	15	1	2011	-	16	1	2011			
21	Аппаратура и оборудование	КЕДР (КСА-Т12 282/РГД 116)									
22	Выбор оборудования	термометр									
23		влагомер									
24											

Рис. 2. Пример отчета

Для автоматизации передачи отчетов и исходных данных в систему в строго определенной форме (регламент именования файлов) решено разработать ПО, частично автоматизирующее этот процесс.

В качестве исходных к этому модулю были выдвинуты следующие требования:

- пользователям должен быть предоставлен простой и удобный web-интерфейс для выполнения вышеописанной задачи;
- система должна предоставлять возможность загружать файлы только определенных форматов;
- при загрузке файлов исследований и отчетов пользователь должен иметь возможность указать ряд параметров (месторождение, куст, скважина, дата исследования, вид исследования), согласно которым производится переименование загружаемых файлов и создание структуры дерева отчетов;
- пользователь за одну загрузку может загружать один файл отчета формата «xlsx», от 1 до 5 файлов формата «las» и от 1 до 5 файлов изображений формата «wvs»;
- при загрузке файлов отчетов в систему, должно формироваться дерево объектов на жестком диске сервера, а так же в служебной таблице БД;
- пользователи должны иметь доступ к дереву отчетов посредством FTP;
- разрабатываемый модуль должен иметь возможность авторизации на уровне БД.

Пример входной формы для загрузки данных разработанного модуля представлен на рис. 3.

Рис. 3. Пример входной формы для загрузки данных

В ходе загрузки данных в систему, периодически возникает необходимость распознать входные данные из выходного формата *.las и загрузить их в БД. Так как стандартного инструмента для этой задачи не существует, было решено разработать инструмент для разбора «las» файлов с последующей загрузкой разобранных данных, через стандартные инструменты СУБД.

Так как «las» формат имеет стандартный ASCII формат, проблем со считыванием файлов не возникло. Однако в зависимости от скважины, наборы данных в различных файлах могут различаться как по количеству, так и по видам исследований. Основная информация о содержании конкретного файла прописана в «шапке» файла. Так как, названия исследований задаются вручную пользователями, от файла к файлу, названия одних и тех же исследований могут варьироваться (сокращения, опечатки). В связи с большим количеством входных файлов (несколько тысяч) исправить названия вручную невозможно. Чтобы решить эту проблему, в разработанной программе была реализована поддержка синонимов.

Дальнейшим шагом при работе с ПГИ является проверка некорректных данных, попавших в буферную зону системы; для автоматизации этого процесса был написан пакет, состоящий из нескольких скриптов на языке PL/SQL, проверяющий входные данные на выходжение их из граничных условий, заданных для каждой категории параметров отдельно.

Мы показали информационный подход, обеспечивающий геологические службы недропользователя актуальными данными ПГИ. Схематично описали процесс сбора, загрузки и верификации данных посредством внедрения данной системы.

Следующим этапом является разработка, на базе представленного «фундамента», различных аналитических алгоритмов, позволяющих решать конкретные прикладные производственные задачи геологических служб недропользователя, основанные на данных ПГИ:

- Подсчет количества скважин, охваченных исследованиями (общий, по каждому виду исследования и по годам), с возможностью верификации;
- Подсчет коэффициентов работающих толщин (max, min, avg) и охвата воздействием с помощью сопоставления данных интерпретации геофизических исследований скважин, интервалов перфорации и данных ПГИ скважин по определению профиля притока/приемистости.
- Сопоставление данных интерпретации геофизических исследований скважин, интервалов перфорации и данных ПГИ скважин по определению текущей нефтенасыщенности.

- Сопоставление данных интервалов перфорации (из базы) и интервалов перфорации промыслово-геофизических исследований скважин по контролю перфорации [3, 4].

Перспективы развития информационной системы

Данные промыслово-геофизических исследований являются хоть и важной, но только частью контента информационной системы, необходимой сотрудникам геологических служб в их деятельности.

В рамках данной работы решается задача создания комплексного информационного хранилища по возможности всех видов исходной геолого-геофизической информации, промысловых данных, результатов проектирования разработки месторождений.

Комплексное хранилище данных корпоративного уровня должно содержать всю информацию по проводимым исследованиям за всю историю работы скважин, поэтому при разработке ХД нужно учесть, что данные по исследованиям имеют разную природу в связи с тем, что проводятся исследования различных типов.

Наибольший эффект можно получить, объединяя все смежные разделы данных, такие как гидродинамические исследования (ГДИС), геолого-технические мероприятия (ГТМ), ПГИ, технологические режимы работы скважин, «шахматки» и т. д. Имея все данные в одном ХД, можно строить сводные и аналитические отчёты любой сложности. Поэтому необходимо было наделять ХД следующими способностями:

- Расширяемость существующей системы новыми разделами данных на базе единого ядра. То есть всегда должна присутствовать возможность добавления нового раздела без изменения структуры базы данных и процедур, обеспечивающих её работу.
- Гибкость структуры. В базе данных необходимо было заложить возможность модификации перечня свойств объектов данных (добавления новых свойств, удаление) без необходимости модификации существующей структуры БД и системы.
- Версионность. Очень полезно знать, в рамках каких организационных процессов появился тот или иной показатель, и из какого источника. Это может помочь пользователю определиться с выбором достоверных показателей.
- Дружественный интерфейс – интерфейс системы должен быть простым и понятным любому пользователю.
- Всем перечисленным требованиям полностью соответствует информационный подход, называемый DataWarehouse:

В системе имеется общее универсальное ядро базы данных;

При хранении данных применяется следующий подход:

- выделяется сущность (пласт, ГИС, ПГИ и т. п.);
- составляется её полное описание: перечисляются все атрибуты (характеристики) сущности;
- данные загружаются и отображаются строго по составленному описанию;
- каждое значение атрибута прикрепляется к производственному процессу, в рамках которого оно получено.

Интерфейс и система представляются отдельными модулями, что позволяет использовать систему в интеграции с любыми программными продуктами.

Концептуальная схема всей системы приведена на рис. 4.

Хранилище данных состоит из 4 уровней:

1. Операционный уровень базы данных. На этом уровне хранятся все исходные данные системы.
2. Уровень обеспечения доступа к данным. На этом уровне находится набор процедур для извлечения/преобразования/загрузки. Является промежуточным уровнем между операционным уровнем и информационным.
3. Уровень метаданных. На этом уровне находятся словари данных и описание сущностей.
4. Информационный уровень. На этом уровне находятся данные для построения отчётов и анализа.



Рис. 4. Концептуальная схема системы

В разработанной системе операционный уровень состоит из таблиц объектов, фактов, и процессов, в ходе которых получены все данные. Каждый объект может содержать в себе сколько угодно разнородных многомерных таблиц, состоящих из наборов значений, каждое из которых указывает на процесс, который послужил причиной появления этого значения в хранилище данных. Также каждый объект может быть связан с другим объектом (по правилу, описанному в каталоге).

Уровень метаданных в разработанной системе представляется как каталог данных, который является дескриптором всех возможных ПГИ сущностей и словарь значений (а также набор синонимов и словоформ для каждой словарной лексемы для поиска и загрузки данных). Каталог описывает правила, по которым можно объединять между собой объекты различных классов, свойства, которые имеет каждый класс, их тип данных, обязательность и т. п.

Уровень обеспечения доступа к данным реализован набором процедур, которые собирают и через определенное время обновляют представлением всех ссылок на значения в кубе. По каждой сущности составляется своё представление из таблицы изменений и представление на основе ядра, которое вместе с представлениями остальных сущностей образует «звезду» OLAP'а. Но так как все объекты связаны между собой, образуется «паутина» из всех представлений объектов, в центре которой находится ядро ХД. Представления собираются на основе каталога данных, и при каждом его изменении меняется вся паутина ХД. Каждое значение любого объекта получено в ходе какого-то процесса и хранит ссылку на обоснование на проведение этого мероприятия.

В ходе дальнейшей работы планируется приступить к автоматизации ранее описанных аналитических задач работы с ПГИ.

Выводы

Рассмотрена возможность применения информационно-технологического подхода к задаче систематизации и обработки данных промыслово-геофизических исследований с целью анализа выработки запасов и вопрос автоматизации этой задачи. Формализован этап поступления исходных данных и отчетной документации в хранилище данных:

- создан регламент для работы с подрядчиками;
- создан ряд автоматизированных шаблонов отчетной документации;
- создано web-приложение загрузки поступающих данных от подрядчика в систему;
- реализован инструмент автоматического разбора исходных «las» файлов и загрузки данных в систему;
- реализован ряд скриптов проверки корректности входных данных.

СПИСОК ЛИТЕРАТУРЫ

1. Ошлакова А.С. Промыслово-геофизические исследования скважин при контроле за разработкой // IV регион. научно-техн. конф. молодых специалистов ОАО «ТомскНИПИнефть»: Сб. тез. докл. – Томск: «ТомскНИПИнефть», 2011. – 232 с., С. 80–85.
2. Шагиев Р.Г. Методические указания по комплексированию и этапности выполнения геофизических, гидродинамических и геохимических исследований нефтяных и нефтегазовых месторождений. – М.: Экспертнефтегаз, 2002. – 15 с.
3. Хуснулин М.Х. Геофизические методы контроля разработки нефтяных пластов. – М.: Недра, 1989. – 123 с.
4. Ипатов А.И., Кременецкий М.И. Геофизический и гидродинамический контроль разработки месторождений углеводородов. – М.: РХД, 2006. – 45 с.

Поступила 26.09.2010 г.