

*На правах рукописи*



**ЩЕРБИЧ Алексей Юрьевич**

**МАТЕМАТИЧЕСКИЕ МОДЕЛИ  
И МЕТОДЫ КОНТРОЛЯ КАЧЕСТВА  
МЕТАДАНЫХ СЕЙСМОРАЗВЕДКИ  
ДЛЯ ФОРМИРОВАНИЯ БАНКОВ  
ГЕОЛОГО-ГЕОФИЗИЧЕСКИХ ДАННЫХ**

**Специальность 05.13.18 – Математическое моделирование,  
численные методы и комплексы программ**

**АВТОРЕФЕРАТ**  
**диссертации на соискание ученой степени**  
**кандидата технических наук**

**Тюмень – 2016**

Работа выполнена на кафедре алгебры и математической логики Федерального государственного автономного образовательного учреждения высшего образования «Тюменский государственный университет»

**Научный руководитель:** доктор физико-математических наук, профессор  
**Кутрунов Владимир Николаевич**

**Официальные оппоненты:** **Туренко Сергей Константинович**,  
доктор технических наук, профессор,  
Федеральное государственное бюджетное  
образовательное учреждение «Тюменский  
индустриальный университет», заведующий  
кафедры прикладной геофизики

**Алтунин Александр Евгеньевич**,  
кандидат технических наук, ООО «Тюменский  
нефтяной научный центр», старший эксперт  
экспертного совета

**Ведущая организация:** Филиал ООО «ЛУКОЙЛ-Инжиниринг»  
«КогалымНИПИнефть» в г. Тюмени

Защита диссертации состоится 15 декабря 2016 г. в 14-00 часов на заседании диссертационного совета Д 212.274.14 при ФГАОУ ВО «Тюменский государственный университет» по адресу 625003, г. Тюмень, ул. Перекопская, 15А, ауд. 410.

С диссертацией можно ознакомиться в библиотеке ФГАОУ ВО «Тюменский государственный университет» и на сайте <https://diss.utmn.ru/sovet/diss-sovet-212-274-14/zashchita/270504/>.

Автореферат разослан « \_\_\_\_\_ » \_\_\_\_\_ 2016 г.

*Ученый секретарь  
диссертационного совета*



*Е. А. Оленников*

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы исследований.** На сегодняшний день в компаниях, занимающихся разведкой и разработкой месторождений полезных ископаемых, накопленные объемы данных варьируются в среднем от нескольких десятков терабайт до нескольких петабайт и более в цифровом исчислении. При этом большую часть геолого-геофизических материалов занимают данные сейсмической разведки. Преследуя цель иметь комплексную геолого-технологическую модель, организации зачастую осуществляют сбор информации в избыточном количестве, отодвигая вопросы контроля качества таких данных на долгую перспективу. Вопросы управления большими объемами цифрового и вещественного материала, в свою очередь, обозначают следующие актуальные проблемы геолого-геофизической отрасли:

- Как правило, цифровые и вещественные наборы данных внутри больших организаций хранятся в разрозненных формах, множественных хранилищах, часто не имеющих ни логической, ни физической связей.

- Отсутствие комплексной модели хранения данных вынуждает пользователей тратить существенное время на поиск информации по неформализованным критериям, при этом результат их поисков не всегда достоверен в силу физической недоступности всего спектра данных и ряда других технических проблем.

- Отсутствие централизованного репозитория с логически выверенной структурой является причиной дублирования данных, что накладывает на компанию, в том числе, и коммерческие издержки по их хранению и приобретению.

- Без стандартов оценки качества материала часть данных в результате оказывается в таком состоянии, из которого невозможно извлечь практическую пользу.

- Формирование банков данных – это сложный, зачастую многолетний процесс. Принятие решений в пользу создания информационной модели порой воспринимается как серьезный риск, который осложнит возможность управления уже устоявшимися бизнес-процессами.

- Является проблемой отслеживание всей истории изменений по мере использования данных.

- Одна из главных проблем – это отсутствие методик как в части системы организации всего спектра геолого-геофизических данных, так и в части их верификации при наполнении банка данных и по мере дальнейшего использования.

Опираясь на практический опыт по созданию банков данных геолого-геофизических материалов, автор склонен утверждать, что промышленная польза от наличия информационных репозиториях в нефтегазовых компаниях

не может быть недооценена. Видимый невооруженным глазом процент брака среди данных, с которыми работает типичный пользователь (геолог, геофизик, геодезист, оператор и др.) может и должен быть исключен средствами автоматизации. Современный технологичный мир с высоким взаимным проникновением математики, геологии, геофизики позволяет оптимизировать значительные объемы ручного труда, используя при этом абсолютно несложные и недорогие с точки зрения реализации методы.

Стратегическая задача, которую сегодня ставят перед собой многие ресурсодобывающие компании, заключается в повышении нефтеотдачи за счет внедрения современных информационных решений в области анализа уже имеющихся данных. Приоритет задач контроля качества технологических процессов становится все выше с ростом объемов и частоты поступления новых данных. В связи с обозначенными причинами необходимость формирования информационных ресурсных моделей и разработка стандартов оценки качества данных подтверждают **актуальность темы** диссертационной работы.

**Целью диссертационной работы** является разработка средствами математического моделирования методической основы для оценки качества метаинформации в структуре сейсмических данных и реализация численных методов контроля качества в программных комплексах, применяемых в процессе формирования и сопровождения геолого-геофизических банков данных.

#### **Задачи работы:**

1. Провести обзор литературных источников, исследующих вопросы управления информацией и контроля качества данных в нефтегазовой отрасли.
2. Провести анализ практических способов управления данными разведки и разработки в отечественной и зарубежной практике.
3. Рассмотреть математические модели, позволяющие вести контроль качества технологических процессов.
4. Провести систематизацию данных сейсморазведки и выделить множества метаданных, подлежащих анализу, на предмет контроля качества.
5. Представить концептуальную математическую модель для количественной оценки качества данных сейсморазведки на основе анализа метаданных.
6. Представить математические модели авторских численных методов, использовавшихся в реальных практических задачах для контроля качества сейсмических данных.
7. Описать функциональную архитектуру программных комплексов, в которых реализованы авторские численные методы.
8. Продемонстрировать полезность авторских алгоритмов и методик, внедренных в реальные производственные процессы ведущих российских и зарубежных нефтегазовых компаний.

**Объектом исследования** являются данные сейсмической разведки и сопутствующие им материалы и атрибуты сейсмической съемки.

**Предмет исследования** – средства анализа для наполнения и оценки качества геолого-геофизических банков данных.

**Методами исследования** являются методы математического моделирования, методы теории принятия решений, методы нечеткой логики, анализ производственных процессов оценки качества данных сейсмической разведки, аналитические и численные алгоритмы, расчеты и их программная реализация, анализ эффективности полученных результатов.

**На защиту выносятся** следующие результаты, соответствующие четырем пунктам паспорта специальности 05.13.18 – математическое моделирование, численные методы и комплексы программ по техническим наукам:

*Пункт 1. Разработка новых математических методов моделирования объектов и явлений.*

1. Предложенная автором концептуальная двухуровневая математическая модель для анализа качества сейсмических данных.

*Пункт 3. Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий.*

2. Разработанные автором численные методы, использующиеся для повышения эффективности контроля качества и загрузки атрибутов сейсмических данных:

- Метод верификации форматов данных сейсморазведки, поступающих на магнитных лентах, работающий по алгоритму дерева принятия решений;
- Усовершенствованный алгоритм Джарвиса для расчета выпуклого контура 2D и 3D сейсмической площади;
- Сформулированный математически и разработанный вычислительный метод контроля качества геометрии 2D сейсмического профиля. Отдельные блоки метода были известны ранее, выполнялись специалистами вручную и не связывались между собой в единый расчетный комплекс, поэтому ранее были глубоко эмпирическими, присущими конкретным исполнителям;
- Метод контроля корректности и автоматического исправления текстовых и бинарных атрибутов в служебных заголовках файлов SEG Y, основанный на нейронной сети Хэмминга. Применительно к метаданным сейсморазведки метод реализован впервые.

*Пункт 4. Реализация эффективных численных методов и алгоритмов в виде комплексов программно-ориентированных программ для проведения вычислительного эксперимента.*

3. Разработанные автором программные комплексы SeismicLoader, GeoStore и SmartSEG Y. Указанные программные комплексы реализуют предложенный спектр алгоритмов и используются в коммерческих и государст-

венных нефтегазовых компаниях для вычислительных экспериментов с целью оценки качества атрибутов сейсморазведки и их последующей загрузки в банки данных. Программный комплекс SmartSEG Y и программный модуль построения невыпуклых оболочек ConcaveHullBuilder прошли государственную регистрацию под номерами 2015614190 и 2016610147 соответственно. Неполный список компаний, использующих разработанное ПО:

- Total – французская нефтяная компания, ведущая общемировой банк данных сейсморазведочной информации;
- ReMASA – компания, являющаяся официальным оператором национального геолого-геофизического банка данных Аргентины;
- РТТЕР – национальная компания Таиланда, ведущая банк данных сейсмике по всем разведочным активам компании, находящимся в разных участках земного шара;
- Лукойл – крупнейшая коммерческая российская нефтяная компания, ведущая банк данных геолого-геофизической информации по Западной Сибири;
- Южморгеология – национальная российская компания в структуре Росгеолфонда, ведущая учет сейсморазведочной информации на шельфе Черного моря.

*Пункт 5. Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента.*

4. Автором предложены комплексные методы оценки качества и актуальности как поступающей, так и хранимой в БД сейсморазведочной информации. Алгоритмы внедрены автором как в собственные программные комплексы, так и в функциональную часть программного комплекса Decision Space Data Quality компании Halliburton и используются в сотнях компаний по всему миру. Указанный программный комплекс используется также и в родственных областях исследований: ГИС, в анализе оперативных данных реального времени в процессе бурения, в анализе информации по добыче.

Таким образом, в соответствии с формулой специальности 05.13.18, в диссертации представлены оригинальные результаты одновременно из трех областей: математического моделирования, численных методов и комплексов программ.

**Научная новизна основных результатов**, полученных в ходе исследования, по трем областям специальности 05.13.18, заключается в следующих положениях:

*Математическое моделирование*

1. Предпринята попытка повышения качества сейсмических данных за счет служебной информации (атрибутов метаданных), а не сейсмического сигнала как такового.

2. Новыми являются предложенные алгоритмы оценки качества геометрии сейсмического профиля; ранее для оценки качества геометрии применялся только визуальный экспертный контроль.

3. Показана возможность использования известных математических моделей для количественной оценки качества информации, поступающей на загрузку в геолого-геофизический банк данных. Эта оценка в ряде случаев может служить «маркером» для принятия решения о допустимости дальнейшей работы с конкретным набором данных.

4. Принципиально новой является концепция отображения численных методов и используемого ими потока атрибутов сейсмоданных на двухуровневую математическую модель с целью принятия решения о численной оценке качества.

5. Благодаря предложенной двухуровневой математической модели удалось формализовать большой объем разрозненной информации, обладающей полезными для анализа свойствами. Математическая модель оценки качества определила универсальный механизм работы авторских численных методов, реализованных в программных продуктах. Таким образом, роль человеческого фактора в анализе качества данных сведена к минимуму.

#### *Численные методы*

6. Алгоритм построения выпукло-вогнутых контуров сейсмических площадей является полностью авторским. В его основу легла методика расчета выпуклой оболочки алгоритмом Джарвиса. Далее была сформулирована и доказана теорема о возможности формирования вогнутой оболочки на основе выпуклой и с учетом входных условий предложен метод преобразования выпуклой оболочки в вогнутую. Данный результат получил широкое аналитическое и коммерческое применение, в частности, в перечисленных выше компаниях, и в значительной степени упростил процесс формирования карт сейсмо-разведочных работ.

7. Предложен эффективный способ фильтрации сейсмических трасс по пространственному критерию, основанный на переносе метода трассировки луча из вычислительной геометрии в анализ данных сейсморазведки. В программном комплексе PetroVision данное решение позволило быстро реализовать отбор данных с карты в соответствии с территориальными правами доступа. Ранее в подобных ГИС-ориентированных системах использовались визуальные, а не вычислительные средства отбора данных, поэтому алгоритм является очевидным новшеством.

8. Алгоритм мониторинга и последующего исправления ошибок в атрибутах служебных заголовков файлов сейсмоданных на основе нейронной сети Хэмминга является нововведением в части автоматизированного наполнения банка данных и автокоррекции загружаемых атрибутов.

9. Разработанные средства оценки и корректировки сейсморазведочной информации, хранящейся в банке данных, являются универсальными, поэтому их методическая значимость состоит в том, что они легко проецируемы на родственные области исследований: анализ данных ГИС, анализ атрибутов добычи и пр.

10. Подготовлена концептуальная методическая база знаний оценки качества атрибутов сейсмической съемки, представленная основными алгоритмами. Некоторые их алгоритмов опираются в своей основе на известные математические модели, однако применение этих моделей для исследования проблем контроля качества метаданных сейсморазведки является новым авторским подходом.

#### *Комплексы программ*

Спроектированы и разработаны программные комплексы, в которых реализованы все вышеобозначенные научные методы. Гибкость архитектуры авторских программных комплексов SeismicLoader, GeoStore, SmartSEGy позволяет использовать существующие методы на широком спектре данных сейсмической разведки и, кроме того, дополнять его новыми простейшими методами оценки качества без внесения изменений в исходный код. Программные комплексы были реализованы автором для непосредственного их использования в нефтегазовых компаниях и государственных структурах, ведущих централизованный сбор геолого-геофизических данных. Программные продукты SeismicLoader и GeoStore были разработаны автором в процессе работы в ООО «Геолідер», внедрение в производство осуществлялось непосредственно автором. Комплекс SmartSEGy разработан автором в процессе работы в подразделении Landmark нефтесервисной компании Halliburton. Здесь же автором выполнена реализация логики принятия решения концептуальной математической модели в функциональной части программного комплекса Decision Space Data Quality.

**Достоверность и обоснованность** полученных результатов подтверждаются экспертными оценками пользователей, применяющих реализованные методы в реальном производстве. Кроме того, в основу формирования базы методик, разработанных для оценки качества, легли реальные проблемы, «тормозящие» производственные процессы в нефтегазовых компаниях. В процессе тестирования разработанных методов результаты сопоставлялись с ранее полученными экспертными заключениями. Быстрое распространение и широкое применение авторских комплексов SeismicLoader, GeoStore, SmartSEGy в крупных нефтегазовых компаниях также говорит о соответствии результатов его работы реальным практическим ожиданиям.

**Практическая значимость** полученных результатов состоит в применении программных комплексов SeismicLoader, GeoStore, SmartSEGy для кон-

троля качества служебных данных сейсморазведки, поступающих на загрузку в геолого-геофизические банки данных. Методики и алгоритмы разрабатывались прежде всего для практического применения в компаниях, занимающихся формированием банков данных разведки и разработки месторождений полезных ископаемых. Укажем лишь на некоторые из областей применения полученных результатов в реальной производственной практике:

- автоматизация процессов загрузки атрибутов сейсмической разведки в банк данных;
- исправление ошибок в метаданных с минимизацией человеческого фактора;
- построение карт сейсмических съемок;
- комплексная оценка качества поступающей и хранящейся сейсмической информации;
- отбор данных на основе критериев, согласованных с корпоративными регламентами доступа к информации.

**Апробация результатов.** Основные положения и результаты диссертации докладывались и обсуждались на следующих научно-практических мероприятиях:

1. III Международный Симпозиум пользователей программного обеспечения PetroVision (По, Франция, 2007 г.).
2. Защита концепции внедрения национального сейсмического Банка Данных Таиланда в компании РТТЕР (Бангкок, Таиланд, 2010).
3. Вторая научно-практическая конференция «Новые технологии поиска нефти и газа» (Тюмень 2014).
4. Общемировой инновационный форум пользователей программного обеспечения Landmark (Париж, Франция, 2014).
5. Международная научно-практическая конференция «NDR 2014» (Баку, Азербайджан, 2014).
6. 70-я международная научно-практическая конференция «Нефть и газ – 2016» (Москва, РГУ нефти и газа имени И.М. Губкина, 2016).
7. Научно-техническая конференция «Современные технологии нефтегазовой геофизики» (Тюмень, ФГБОУ ВО «Тюменский индустриальный университет», 2016).
8. Научные семинары кафедры алгебры и математической логики Тюменского государственного университета.

**Публикации.** Основные результаты диссертации опубликованы в 10 печатных работах, в том числе 2 статьи в рецензируемых изданиях, рекомендованных ВАК для предоставления основных научных результатов диссертаций на соискание ученой степени доктора или кандидата наук. Получено 2 свидетельства об официальной регистрации программ для ЭВМ.

**Личный вклад.** Результаты, составляющие основное содержание диссертации, получены автором самостоятельно. Авторскими являются как численные алгоритмы, так и их реализация в программных комплексах. Совместно с научным руководителем была разработана концептуальная математическая модель и универсальный механизм принятия решения о качестве данных, реализованный с помощью предложенных автором численных методов.

**Структура и объем диссертации.** Диссертация состоит из введения, четырех глав, заключения, списка сокращений и условных обозначений, списка литературы, включающего 134 наименования, и четырех приложений. Диссертация изложена на 164 страницах машинописного текста, включая 55 рисунков и 10 таблиц.

## СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во **Введении** описывается круг проблем области исследования. На основе требований современной геолого-геофизической отрасли приводится ряд примеров, где существует необходимость разработки стандартов оценки качества, методов математического моделирования, применения численных методов для повышения эффективности управления данными сейсмической разведки. Учитывая показанные факты, ставятся цели и задачи исследования, демонстрируется научная и практическая значимость работы.

**Глава 1** носит обзорный характер. В самом ее начале приведен анализ библиографического материала, освещающего проблемы управления данными сейсморазведки. На основе анализа отечественной и зарубежной литературы приводится классификация методов контроля качества в области управления данными. Рассмотрены работы Д. Салливана, публикации международных сообществ DAMA, Energetics, EAGE. Проведен обзор математических моделей принятия решений, связанных с контролем качества производственных процессов. Рассмотрены работы авторов В. И. Бодрова, Т. Я. Лазаревой, Ю.Ф. Мартемьянова, Г. С. Альтшуллера, Б. Г. Литвака, В.С. Василенка, В.А. Глезер, Е.А. Грампа и др. Анализируются положения теории принятия решений, особенно в части управления качеством производства. Далее вводятся определения понятий «банк данных», «модель данных», сформулированные на основе определений ГОСТ и некоторых технических энциклопедий. Автор приводит описание функциональных компонентов геолого-геофизического банка данных, а также обзор российских и зарубежных программных продуктов, содержащих эти компоненты. В обзоре рассмотрены программные продукты ProSource и InnerLogix компании Schlumberger, PetroBank и DSDQ компании Halliburton, PetroVision компании Геолидер, Whereoil компании KADME, а также некоторые решения компаний СибГеоПроект, Баспро, СибНАЦ. Банк данных геолого-геофизической информации рассмот-

рен и с позиции особенностей его программно-аппаратной архитектуры. Показаны современные решения по управлению данными сейсморазведки и навигации на магнитных лентах, в промышленных форматах записи SEG Y, SEG D, UKOOA, SPS. Предложены для рассмотрения некоторые варианты современных ленточных библиотек компаний IBM, HP, Quantum.

Анализ предметной области указывает на необходимость некоторого упорядочения элементов с выделением крупных блоков объект – математическая модель – решение – результат. В каждом из показанных элементов комплексной инфраструктуры выделяются проблемные компоненты.

В **Главе 2** производится формализация данных сейсмической разведки, их структуризация, и описывается процесс построения двухуровневой математической модели, целью которой является анализ качества данных сейсморазведки как на уровне отдельных вычислительных методов, так и на уровне комплексной количественной оценки качества.

В первую очередь приводится классификация данных сейсморазведки и выделение групп для проведения технологического контроля качества (рис. 1).

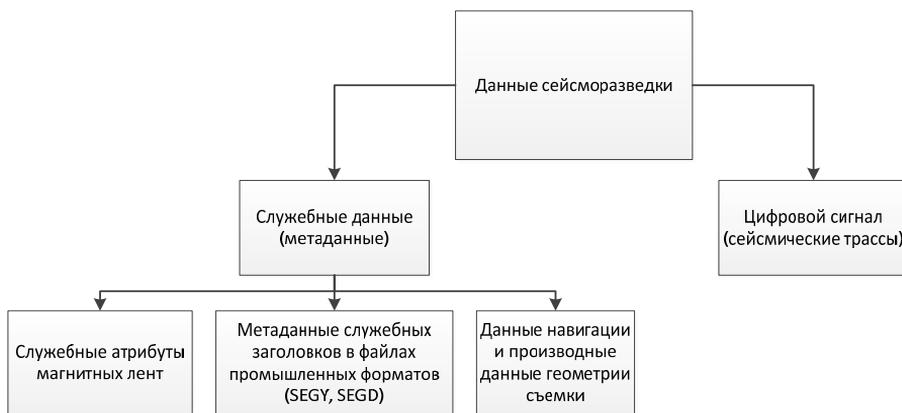


Рис. 1. Иерархия данных сейсморазведки в контексте исследования

Предлагаемая автором концепция контроля качества предполагает проведение технологического контроля не цифрового сейсмического сигнала, а служебной информации, то есть метаданных. Учитывая это условие, будем использовать термины «анализ качества данных» и «контроль качества данных» применимо к метаданным, а не к цифровому сигналу.

В структуре сейсмической информации трем выделенным группам метаданных (рис. 1) соответствует три множества методов контроля качества:

$T = \{M_1^T, M_2^T, \dots, M_n^T\}$  – множество методов оценки качества данных на магнитных лентах, где  $n$  – количество применяемых для контроля качества методов;

$D = \{M_1^D, M_2^D, \dots, M_p^D\}$  – множество методов оценки качества дисковых сейсмических данных, где  $p$  – количество применяемых для контроля качества методов;

$N = \{M_1^N, M_2^N, \dots, M_k^N\}$  – множество методов оценки качества навигационных данных, где  $k$  – количество применяемых для контроля качества методов.

*Процесс принятия решения* о качестве данных  $j$ -й группы, где  $j \in \{T, D, N\}$ , будем понимать как последовательность применения методов соответствующего  $j$ -го множества. Введем обозначения процесса контроля качества данных  $j$ -й группы:

$$P^j = \{M_1^j, M_2^j, \dots, M_l^j\}, j \in \{T, D, N\}, l \in \{n, p, k\} \quad (1)$$

$j$ -ю группу данных обозначим как  $S^j$ .

Каждый из методов  $M_i^j$  процесса  $P^j$  отвечает за анализ некоторой структурной области  $S_i^j \in S^j$ . Иными словами, некоторой исследуемой области данных  $S_i^j$  ставится в соответствие метод  $M_i^j$ .

Экспертно назначается вес  $W_i^j$  – уровень значимости метода  $M_i^j$  в процессе  $P^j$ . Наложим на веса условие:

$$\sum_{i=1}^n W_i^j = 1 \quad (2)$$

**Определение 1** *Предельным уровнем качества области данных  $S_i^j$ , исследуемой методом  $M_i^j$ , называется численная характеристика максимального ожидаемого качества этой выделенной области.*

Таким образом, имеет место функция:

$$Fmax_i^j: S_i^j \xrightarrow{M_i^j} Qmax_i^j, i = 1..l, l \in \{n, p, k\}, j \in \{T, D, N\} \quad (3)$$

**Определение 2** *Рассчитанным уровнем качества области данных  $S_i^j$ , исследуемой методом  $M_i^j$ , называется численная характеристика реально рассчитанного качества этой выделенной области.*

$$Fрасч_i^j: S_i^j \xrightarrow{M_i^j} Qрасч_i^j, i = 1..l, l \in \{n, p, k\}, j \in \{T, D, N\} \quad (4)$$

**Определение 3** *Нормированным уровнем качества области данных  $S_i^j$ , оцениваемой методом  $M_i^j$ , назовем функцию  $F_i^j: S_i^j \rightarrow [0,1]$ . Имея характеристики предельного и рассчитанного уровней качества, можно задать уравнение функции  $F_i^j$ :*

$$Q_{\text{норм}}^j = F_i^j(S_i^j) = \frac{Q_{\text{расч}}^j}{Q_{\text{max}}^j} \quad (5)$$

Имея параметры оценки качества выделенных областей данных  $S_i^j$ , можно определить нормированную оценку качества данных, исследуемых конкретным процессом  $P^j$ .

**Определение 4** Нормированной оценкой качества данных  $j$ -й группы, исследуемой процессом  $P^j$ , будем называть линейную комбинацию нормированных уровней качества, рассчитанных используемыми в процессе  $P^j$  методами  $M_i^j$ , т. е.

$F^{P^j}: P^j \rightarrow [0,1]$  – функция, определяющая нормированную оценку качества данных по процессу  $P^j$ .

По определению,

$$F^{P^j} = W_1^j * F_1^j + W_2^j * F_2^j + \dots + W_l^j * F_l^j = \sum_{i=1}^l W_i^j * F_i^j \quad (6)$$

Таким образом, *первый уровень концептуальной математической модели* определяет логику расчета количественной оценки качества данных отдельно взятой группы. Представленная математическая модель является примером практического применения классической модели расчета интегральной оценки качества производственных процессов из теории принятия решений. В данном случае интегральной оценкой является нормированная оценка качества данных сейсморазведки. Представленную математическую модель принятия решения можно графически изобразить в виде схемы (рис. 2):

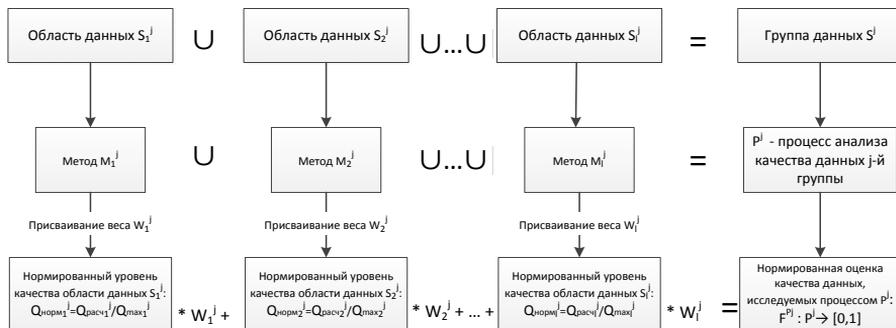


Рис. 2. Модель принятия решения о качестве данных  $j$ -й группы

Процесс  $P^j$  представляется графически в виде графа предшествования  $G^{P^j} = \langle F_i^j, M_i^j \rangle$ , где вершинами являются нормированные уровни качества

$F_i^j$  соответствующих областей данных  $S_i^j$ , а ребрами – методы  $M_i^j$  с присвоенными весами  $W_i^j$  (рис. 3):

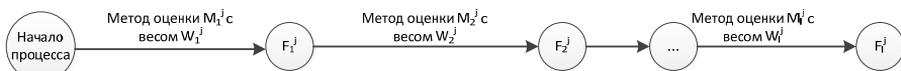


Рис. 3. Модель процесса оценки качества данных  $j$ -й группы в виде графа предшествования

На *втором уровне концептуальной математической модели* обращается внимание на логику принятия решения, реализуемую отдельно взятым методом  $M_i^j$ . Обратим внимание на то, что каждый из используемых в конкретном процессе  $P^j$  методов  $M_i^j$  является вычислительным алгоритмом, работающим в соответствии с собственной математической моделью принятия решения. В общем виде математическая модель метода контроля качества представляется в виде кортежа:

$$M_i^j = \langle A_i^j, L_i^j, Qmax_i^j, LQнорм_i^j, W_i^j \rangle, \quad (7)$$

где  $A_i^j$  – множество атрибутов, задающих область  $S_i^j$ ;  $L_i^j$  – численный алгоритм, устанавливающий наличие брака («некачественных» данных) среди атрибутов множества  $A_i^j$ ;  $Qmax_i^j$  – предельный уровень качества области  $S_i^j$ ;  $LQнорм_i^j$  – численный алгоритм расчета нормированной оценки качества области  $S_i^j$ ;  $W_i^j$  – нормированный вес метода  $M_i^j$  в процессе  $P^j$ .

В ходе практической работы методы контроля качества разрабатывались автором таким образом, чтобы множества атрибутов  $A_i^j$  максимально соответствовали множествам метаданных, загружаемым в результате в геолого-геофизический банк данных. Этим обеспечивается качество служебной информации, хранимой в банке данных. Автор обращает внимание на особенности отображения атрибутов метаданных сейсмоки на логическую модель данных POSC Epicenter.

В **Главе 3** описываются разработанные автором методы контроля качества для каждой из трех групп метаданных сейсмоки. Всего автором предлагается 9 методов, особенности каждого из которых рассматриваются в соответствии со структурой, определенной в (7). Сводный перечень методов контроля качества представлен в таблице 1.

Из множества разработанных и представленных в табл. 1 методов покажем в автореферате лишь некоторые численные алгоритмы  $L_i^j$ , входящие в структуру моделей (7).

## Перечень разработанных методов контроля качества

	Метаданные сейсмике на магнитных лентах	Метаданные промышленных форматов SEG Y/SEG D	Навигационные данные
Метод контроля физического состояния ленты	●		
Метод верификации формата	●		
Метод контроля полноты атрибутивного состава	●	●	
Метод оценки качества инкапсуляции данных	●		
Метод восстановления атрибутов служебных заголовков на основе нейросети Хэмминга		●	●
Усовершенствованный алгоритм Джарвиса для расчета невыпуклого контура сейсмосьемки			●
Метод контроля пропущенных точек сейсмопрофиля			●
Метод контроля кривизны сейсмического профиля			●
Метод отбора пространственных данных с помощью алгоритма трассировки луча			●

**1. Контроль качества метаданных сейсморазведки на магнитных лентах.** В данной группе обратим внимание на метод верификации формата. Задачей метода является идентификация формата записанных на магнитной ленте данных по минимальному количеству первых блоков, а именно – распознавание промышленных форматов SEG Y, SEG D и TAR. Эмпирическим путем было установлено, что исчерпывающий набор атрибутов метаданных содержится максимум в первых трех блоках магнитной ленты, которая содержит данные любого из трех заданных форматов. Автором на основе анализа технической документации составлена генеральная совокупность, состоящая из атрибутов метаданных, необходимых для идентификации форматов (табл. 2). Механизм распознавания формата представим в виде дерева принятия решений, которое построим на основе логики алгоритма ID3 из области интеллектуального анализа данных.

Таблица 2

**Генеральная совокупность атрибутов, необходимых  
для идентификации форматов**

Номер блока данных с ленты ( $Q_1$ )	Размер блока данных (байт) ( $Q_2$ )	Размер предыдущего блока данных (байт) ( $Q_3$ )	Ключевой параметр блока для SEG Y ( $Q_4$ )	Ключевой параметр блока для SEG D ( $Q_5$ )	Ключевой параметр блока для TAR ( $Q_6$ )	Формат
1	Size > 0, Size $\neq$ 400, Size mod 512 $\neq$ 0	NULL	NULL	Fcode $\in$ {0x8015,0x8022, 0x8024,0x8036, 0x8038,0x8042, 0x8044,0x8048, 0x8058,0x8080}	NULL	SEGD
1	Size > 0, Size $\neq$ 400, Size mod 512 $\neq$ 0	NULL	NULL	NULL	NULL	UNKNOWN
1	Size mod 512 = 0	NULL	NULL	NULL	magic = "ustar"	USTAR
1	Size mod 512 = 0	NULL	NULL	NULL	NULL	UNKNOWN
2	Size = 400	3200	Fcode $\in$ [1..5]	NULL	NULL	SEG Y
2	Size = 400	3200	NULL	NULL	NULL	UNKNOWN
2	Size = 400	NULL	NULL	NULL	NULL	UNKNOWN
2	Size > 0, Size $\neq$ 400, Size mod 512 $\neq$ 0	128	NULL	Fcode $\in$ {0x8015,0x8022, 0x8024,0x8036, 0x8038,0x8042, 0x8044,0x8048, 0x8058,0x8080}	NULL	SEGD
2	Size > 0, Size $\neq$ 400, Size mod 512 $\neq$ 0	128	NULL	NULL	NULL	UNKNOWN
2	Size > 0, Size $\neq$ 400, Size mod 512 $\neq$ 0	NULL	NULL	NULL	NULL	UNKNOWN
3	Size = 400	3200	Fcode $\in$ [1;5]	NULL	NUM.	SEG Y
3	Size = 400	3200	NULL	NULL	NULL	UNKNOWN
3	Size = 400	NULL	NULL	NULL	NULL	UNKNOWN

Энтропию исходного множества данных по отношению к атрибуту (свойству)  $Q_i$  будем рассчитывать как

$$H(L, Q_i) = - \sum_{j=1}^{q_i} \frac{m_j}{l} \log_{q_i} \frac{m_j}{l}, \quad (8)$$

где  $L$  – исходное множество данных;  $l$  – количество строк;  $q_i$  – количество различных значений, которые может принимать атрибут  $Q_i$ ;  $m_i$  – количество случаев, в которых атрибут  $Q_i$  принимает очередное значение. Вычислим исходную энтропию множества  $L$  по отношению к целевому свойству «Формат»:

$$H(L, \text{"Формат"}) = -\frac{2}{13} \log_4 \frac{2}{13} - \frac{8}{13} \log_4 \frac{8}{13} - \frac{1}{13} \log_4 \frac{1}{13} - \frac{2}{13} \log_4 \frac{2}{13} \approx 0,773$$

Пусть исходное множество классифицировано посредством атрибута  $Q_i$ . Тогда прирост информации рассчитываем по формуле:

$$\text{Gain}(L, Q_i) = H(L, S) - \sum_{j=1}^{q_i} \frac{|L_j|}{|L|} H(L_j, S), \quad (9)$$

где  $L_j$  – подмножество элементов  $L$ , на которых атрибут  $Q_i$  принимает  $j$ -е значение,  $|L_j|$  – количество элементов в  $L_j$ ,  $|L|$  – количество элементов в  $L$ . Рассчитаем прирост информации для каждого из атрибутов, определяющих множество  $L$ :

$$\begin{aligned} \text{Gain}(L, Q_1) &= 0,773 - \frac{4}{13} * H(\{Q_1 = "1"}, S) - \frac{6}{13} * H(\{Q_1 = 2\}, S) - \frac{3}{13} * H(\{Q_1 = "3"}, S) \\ &= 0,773 - \frac{4}{13} * \left[ -\frac{1}{4} * \log_4 \frac{1}{4} - \frac{2}{4} * \log_4 \frac{2}{4} - \frac{1}{4} * \log_4 \frac{1}{4} \right] - \frac{6}{13} * \left[ -\frac{1}{6} * \log_4 \frac{1}{6} - \frac{4}{6} * \log_4 \frac{4}{6} - \frac{1}{6} * \log_4 \frac{1}{6} \right] - \frac{3}{13} * \left[ -\frac{1}{2} * \log_4 \frac{1}{2} - \frac{1}{2} * \log_4 \frac{1}{2} \right] \approx 0.138; \end{aligned}$$

...

$$\text{Gain}(L, Q_4) = 0,773 - \frac{11}{13} * H(\{Q_4 = "NULL"}, S) - \frac{2}{13} * H(\{Q_4 = [Fcode = 1.5]\}, S) \approx \mathbf{0.3097};$$

...

$$\text{Gain}(L, Q_6) = 0,773 - \frac{12}{13} * H(\{Q_6 = "NULL"}, S) - \frac{1}{13} * H(\{Q_6 = [magic = "ustar"]\}, S) \approx 0.196.$$

Максимальный прирост информации обнаруживается при классификации по атрибуту  $Q_4$ , который является корневым элементов дерева решений.

Проводя далее аналогичные расчеты, формируем конечный вид дерева решений (см. рис. 4).

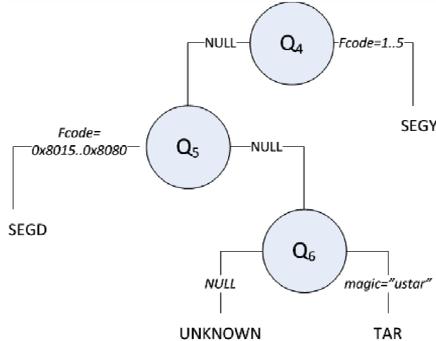


Рис. 4. Дерево решений для распознавания формата сейсмоданных на магнитной ленте

Использованная для расчета дерева математическая модель позволила установить, что атрибуты  $Q_1, Q_2$  и  $Q_3$  являются незначимыми в принятии решения о формате сейсмоданных. Нормированный уровень качества для описанного метода может принимать 2 возможных значения:  $Q_{\text{норм}} \in \{0,1\}$ , где 0 означает, что формат не распознан, а 1 – признак распознанного формата.

**2. Контроль качества метаданных файлов промышленных форматов SEGY.** В этой группе метаданных представим метод восстановления атрибутов служебных заголовков. Метод решает задачу корректировки текстовых и бинарных атрибутов, в которых пользователями были допущены произвольные ошибки в процессе заполнения служебных заголовков SEGY, в особенности – заголовка EBCDIC. Множеством входных данных для метода является множество атрибутов заголовка EBCDIC, участвующих в загрузке в банк данных. Учтем факт, что в банке данных имеется множество словарей, содержащих эталонные значения часто используемых атрибутов метаданных. Так возникает возможность использовать логику ассоциативной памяти, которая может сопоставить искаженное пользователем значение атрибута заголовка EBCDIC с одним из образцовых значений этого атрибута из словаря банка данных. Для решения задачи использована модель на основе искусственной нейронной сети Хэмминга, представленная на рис. 5.

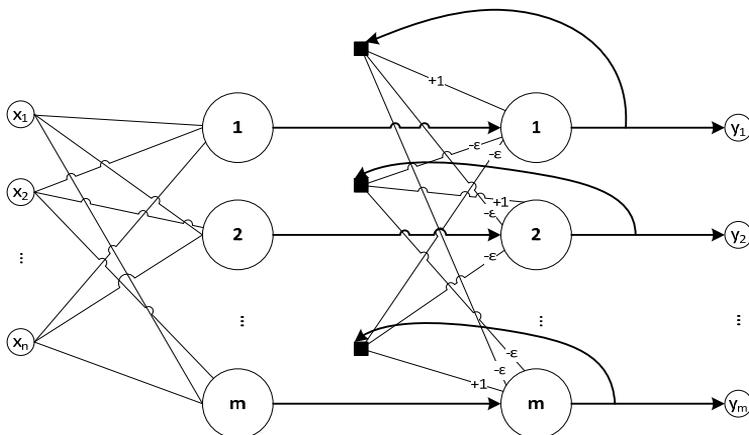


Рис. 5. Структура искусственной нейронной сети Хэмминга

Сеть принимает на вход бинарный вектор  $X$ , кодирующий слово-атрибут, в котором допущена ошибка:  $X = \langle x_1, x_2, \dots, x_n \rangle$ ,  $x_i \in \{-1, 1\}$ . Так как EBCDIC-заголовок SEGY файла использует 8-битную кодировку (EBCDIC или ASCII) для записи атрибутов, то в методе для кодирования были использованы ASCII-коды символов, из которых состоит выбранный атрибут.

Принцип работы сети базируется на идее измерения расстояния Хэмминга от выбранного в заголовке EBCDIC значения атрибута до всех его возможных значений, содержащихся в словаре. Для этого сеть предварительно обучается по следующему закону:

$$w_{ik} = \frac{x_i^k}{2}, i = 0..n - 1, k = 0..m - 1, \quad (10)$$

где  $m$  – количество нейронов первого слоя сети (соответствует числу значений в выбранном словаре);  $w_{ik}$  – весовой коэффициент  $i$ -го синапса  $k$ -го нейрона первого слоя;  $x_i^k$  – это  $i$ -й элемент  $k$ -го образца из словаря.

На вход сети подается вектор  $X$ , в котором закодировано значение выбранного атрибута из заголовка EBCDIC. Далее определяется состояние нейронов первого слоя сети:

$$s_j^{(1)} = \sum_{i=1}^n w_{ij}x_i + T_j, j = 0..m - 1, T_j = \frac{n}{2}, \quad (11)$$

в скобках указан номер слоя.

Для первого слоя значения состояний нейронов одновременно являются значениями их аксонов, то есть  $y_j^{(1)} = s_j^{(1)}$ . Полученные значения  $s_j^{(1)}$  передаются как начальные состояния на аксоны нейронов второго слоя:  $y_j^{(2)} = y_j^{(1)}$ . Расчет нового состояния нейронов второго слоя протекает рекуррентно на основе текущих значений его аксонов. Применяется следующая формула:

$$s_j^{(2)}(p + 1) = y_j^{(2)}(p) - \varepsilon * \sum_{k=0}^{m-1} y_k^{(2)}, k \neq j, j = 0..m - 1, \quad (12)$$

$p$  – номер текущей итерации,  $0 < \varepsilon < \frac{1}{m}$ .

Применяется активационная функция для вычисления новых значений аксонов второго слоя:

$$y_j^{(2)}(p + 1) = f[s_j^{(2)}(p + 1)], j = 0..m - 1. \quad (13)$$

В качестве активационной функции для нейронной сети Хэмминга выбирается функция линейного порога (14):

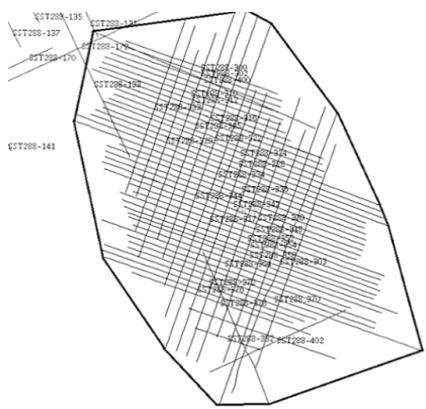
$$f(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < n, \\ n, & x \geq n. \end{cases} \quad (14)$$

Итерационный процесс прекращается, когда на выходе нейронов второго слоя значения перестанут меняться, то есть все значения аксонов, кроме одного, окажутся нулевыми. Порядковый номер нейрона с единственным положительным значением аксона укажет на порядковый номер значения из словаря банка данных, которому максимально соответствует слово заголовка EBCDIC, закодированное во входном векторе  $X$ .

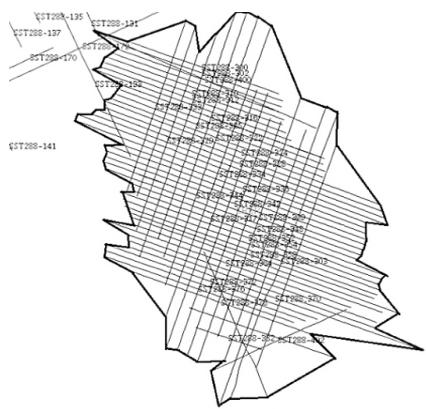
В программном комплексе SmartSEGY реализован алгоритм работы метода, а также возможность в автоматическом режиме исправлять ошибочные атрибуты на те, которые подобраны нейронной сетью. В некоторых случаях, когда сеть не может выдать однозначный ответ (случай совпадения расстояний Хэмминга до двух похожих значений словаря), пользователь-эксперт имеет возможность выбрать наиболее допустимый вариант. Метод существенно сокращает скорость анализа входных данных и упрощает последующие процессы загрузки метаинформации в сейсмический банк данных.

**3. Контроль качества навигационных данных.** Первым из методов контроля качества навигационных данных предлагается авторский метод построения невыпуклых контуров сейсмических съемок. Метод выполняет важную задачу – контроль качества изученности местности методами сейсмической разведки. Пользователь на выходе работы данного метода получает контуры, ограничивающие на карте области сейсмических исследований.

Автор для разработки метода исследует теоретические аспекты построения выпуклых и вогнутых оболочек (рис. 6). Обращается внимание на методику «alpha shapes», методы триангуляции, частные исследования авторов А. Морейра, М. Сантоса, Д.-С. Парк, Се-Лонг О и др., программную библиотеку геометрических методов CGAL.



а) выпуклая оболочка



б) вогнутая оболочка

Рис. 6. Выпуклая и вогнутая оболочки сейсмической съемки

Учитывая особенности геометрии сейсмической съемки, для построения выпуклой оболочки автором вводится допущение, которое предполагает исследование не всего множества точек, образующих линии приема сейсмического сигнала, а лишь крайних точек сейсмических профилей (или инлайн/кросслайнов). Учитывая это допущение, вводятся определения.

**Определение 5.** *Невыпуклой оболочкой* множества точек  $X = \{x_1, x_2, \dots, x_n\}$  будем называть любой многоугольник  $P_1P_2 \dots P_m: P_1 \in X, P_2 \in X, \dots, P_m \in X, m \leq n$ , ограничивающий  $X$ , имеющий при этом меньшую площадь по отношению к выпуклой оболочке для  $X$ .

**Определение 6.** *Входным множеством данных* назовем такое множество  $S = \{[S_{11}S_{12}], [S_{21}S_{22}], \dots, [S_{n1}S_{n2}]\}$ , где  $[S_{i1}S_{i2}]$  – это отрезок с концами в точках  $S_{i1}, S_{i2}$ , а  $n$  – общее число отрезков на плоскости.

Множество точек, составленное из концов отрезков  $[S_{i1}S_{i2}]$ , обозначим за  $S^P$ .

**Определение 7.** *Выпуклая оболочка для  $S$*  – это такое множество  $\text{ConvexHull}(S) = \{[C_{11}C_{12}], [C_{21}C_{22}], \dots, [C_{m1}C_{m2}]\}$ , являющееся выпуклой оболочкой для всех точек  $S_{jk} \in S^P$ , где  $[C_{i1}C_{i2}]$  – это отрезок выпуклой оболочки, ограниченный парой точек из  $S^P$ , а  $m$  – число отрезков выпуклой оболочки  $S, m \leq n$ .

В процессе исследования показывается, что выпукло-вогнутая оболочка может быть рассчитана на основе выпуклой оболочки путем последовательной замены каждого отрезка  $[C_{i1}C_{i2}]$  на два смежных отрезка  $[C_{i1}K]$  и  $[KC_{i2}]$ , где точка  $K \in S^P$ , но не является точкой выпуклой оболочки, то есть лежит внутри нее. Для этого формулируется и доказывается следующая теорема.

**Теорема 1.** Пусть  $S = \{S_1, S_2, \dots, S_n\}$  – конечное множество точек плоскости,  $A = \{A_1, A_2, \dots, A_m\}$  – множество концов отрезков (угловых точек) выпукло-вогнутой оболочки, то есть  $A \subset S$ . Тогда если  $B = \{B_1, B_2, \dots, B_k\}$  – множество концов отрезков выпуклой оболочки, где  $B \subset S$ , то  $B \subset A$ . С учетом доказанной возможности преобразования выпуклой оболочки в вогнутую, итоговый алгоритм разбивается на 2 этапа: построение выпуклой оболочки известным методом Джарвиса и преобразование выпуклой оболочки в вогнутую методом «минимальных углов». Метод минимальных углов итеративно заменяет очередную сторону построенной выпуклой оболочки  $[C_{i1}C_{i2}]$  последовательно на новые отрезки  $[C_{i1}K]$  и  $[KC_{i2}]$ ,  $K \in S^P$ , где  $\angle KC_{i1}C_{i2}$  является минимальным среди всех углов, образованных точками-кандидатами из  $S^P$  и концами отрезка  $[C_{i1}C_{i2}]$ .

Автором проводится анализ работы алгоритма, итоговая вычислительная сложность которого составляет  $O(n^2)$ . Вопросы оптимизации скорости работы алгоритма в данной работе не ставятся.

Для навигационных данных отметим также особенности метода контроля пропущенных точек сейсмического профиля. Задача метода – идентификация сегментов сейсмического профиля, на которых могут отсутствовать точки, его образующие (напр., ПП-пункты приема сигнала). Отсутствие точек на профиле чаще всего проявляется вследствие действий пользователей, редактирующих исходные навигационные файлы сейсморазведки. Иллюстрация работы метода приведена на рис. 7.



Рис. 7. Идентификация пропущенных точек на сейсмическом профиле

Пусть сейсмопрофиль состоит из  $n$  точек, тогда можно выделить  $n - 1$  сегмент профиля. Обозначим последовательно за  $r_1, r_2, \dots, r_{n-1}$  длины сегментов. Тогда

$$r_{\text{ср.}} = \sum_{i=1}^{n-1} r_i / (n - 1) - \text{средняя длина сегмента сейсмопрофиля.}$$

Предполагается, что средняя длина сегмента сейсмопрофиля не должна превышать величины  $r_{\text{ср.}} * k, k \in \mathbb{N}$ , то есть алгоритм контроля качества итеративно проверяет условие:

$$r_i \leq r_{\text{ср.}} * k, k \in \mathbb{N}, i = 1..n - 1 \quad (15)$$

В большинстве случаев  $r_i \rightarrow r_{\text{ср.}}$ , однако введен поправочный натуральный коэффициент  $k$ , определяемый пользователем-экспертом, для учета сложности условий съемки. На практике применялись небольшие значения  $k \in \{1, 2, 3\}$ , в большинстве случаев,  $k = 2$ .

Формула расчета нормированного уровня качества для данного метода выглядит так:

$$Q_{\text{норм}} = 1 - \frac{n_{\text{det}}}{n-1}, \quad (16)$$

где  $n_{\text{det}}$  – число обнаруженных сегментов сейсмопрофиля с отсутствующими на них точками.

В **Главе 4** рассматриваются особенности разработанных автором программных комплексов, в которых реализованы все предлагаемые автором методы контроля качества метаданных. Информационно-аналитическая система предназначена для обеспечения доставки достоверной сейсмической информации в геолого-геофизический банк данных. Некоторые особенности функциональной архитектуры представлены на рис. 8.

В качестве *логической модели банка данных* выбрана промышленная модель POSC Epicenter. Для физической реализации модели использовалась СУБД Oracle.

*Исходными данными*, подлежащими контролю, стали цифровые материалы сейсмической разведки на магнитных лентах (Exbyte, LTO, IBM, DLT) и в файлах стандартизованных форматов (SEG-A,B,C, SEGY, SEG-D, TAR, RODE, TIF, STAGE, UKOOA, SPS).

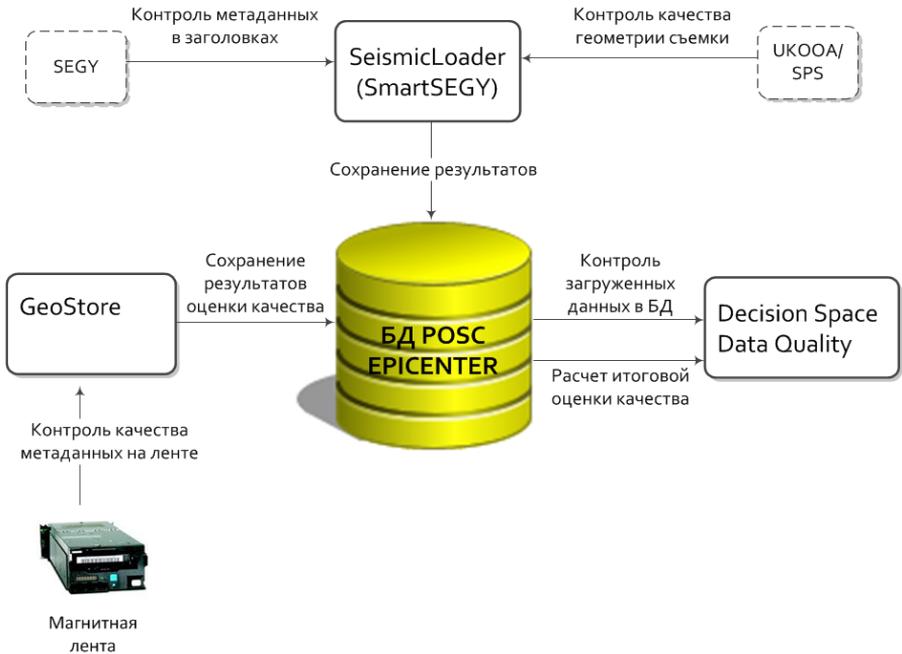


Рис. 8. Функциональная архитектура информационной системы

Приложения *SeismicLoader* и *SmartSEGY* реализуют методы контроля качества метаинформации, содержащейся в файлах стандарта SEGY, и осуществляют контроль качества геометрии сейсмической съемки, описанной в навигационных файлах.

Методы контроля качества метаинформации сейсмических лент реализованы в программном комплексе *GeoStore*. Автор внес собственные расширения в модель данных POSC Epicenter, дополнив ее возможностями учета информации, поступающей с магнитных лент. Таким образом, *GeoStore* обеспечивает контроль качества и загрузку в БД атрибутов сейсмической информации, поступающей на магнитных лентах.

Приложение *Decision Space Data Quality* предоставляет графический интерфейс для наглядного представления результатов расчета итоговой нормированной оценки качества данных. Расчеты нормированной оценки реализованы автором в хранимых процедурах PL/SQL.

Все описанные приложения работают как в полностью автоматизированном режиме для обработки серий однотипных файлов или лент, так и в режиме графического интерфейса с непосредственным участием пользователя.

эксперта. Атрибуты метаданных, прошедшие контроль качества и предусмотренные информационной моделью POSC Epicenter, загружаются в банк данных на долговременное хранение.

Автором на примере некоторых производственных задач наглядно отражены результаты применения предлагаемых методов контроля качества и методики численной оценки качества метаданных.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. На основании анализа задач предметной области проведена систематизация подходов, применимых к контролю качества служебных данных сейсморазведки.

2. Опыт внедрения геолого-геофизических банков данных позволил автору смоделировать типичные рабочие процессы специалистов, осуществляющих первичный контроль качества и загрузку данных сейсморазведки в корпоративные хранилища.

3. Классифицированы цифровые данные по сейсморазведке и выделены 3 множества метаданных сейсмической информации, на которых автором предлагается применять математические методы контроля качества:

- метаданные сейсмической информации на магнитных лентах;
- метаданные служебных заголовков файлов промышленных стандартов хранения сейсмоки;
- метаданные геометрии сейсмической съемки, представленной в навигационных файлах.

4. Представлена авторская концепция двухуровневой математической модели анализа качества данных сейсморазведки:

- *первый уровень математической модели* определяет логику расчета количественной оценки качества данных отдельно взятой группы;
- *второй уровень модели* определяет общие функции и структуру численных методов контроля качества, а также детализирует логику принятия решения о качестве данных отдельно взятыми методами.

5. Разработаны численные методы контроля качества, которые применялись в реальных производственных задачах. В работе представлены математические модели и алгоритмы, использованные для реализации предложенных методов. Всего автором создано 9 численных методов.

6. Разработана информационно-аналитическая система, включающая 3 программных комплекса и отдельные программные модули, в которых реализованы все авторские численные методы контроля качества. Программные комплексы используются для контроля качества и загрузки данных сейсморазведки при наполнении геолого-геофизических банков данных.

7. Продемонстрированы примеры использования авторской методики для входного анализа сейсмического материала.

## СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

### Публикации в рецензируемых научных изданиях, рекомендованных ВАК

1. Щербич А.Ю. Алгоритм расчета выпукло-вогнутого контура площади сейсмической съемки / А.Ю. Щербич, В.Н. Кутрунов // Вестник Тюменского государственного университета. Физико-математические науки, 2014. № 7. – С. 166-177.
2. Щербич А.Ю. Применение математической модели нейронной сети Хэмминга для контроля качества и восстановления некорректных атрибутов метаданных из заголовков сейсмических файлов / А.Ю. Щербич, В.Н. Кутрунов // Вестник Иркутского государственного технического университета. 2016. № 2. – С. 50-60.

### Публикации в трудах международных конференций

3. Scherbich A. Data federation as a way of NDR data sources integration and QC [Электронный ресурс] / A. Scherbich // Proceedings of “NDR2014” conference. – Baku, 2014. – URL: <http://www.energistics.org/calendar/2014/9/29/ndr2014> (дата обращения: 29.02.2016).
4. Щербич А.Ю. Реализация алгоритма фильтрации случайных шумов на сейсмических трассах / А.Ю. Щербич // Математические методы в технике и технологиях (ММТТ-20): сб. трудов XX Междунар. научн. конфер.Т. 6. Секц. 12. – Ярославль: Изд-во Ярослав. гос. техн. ун-та, 2007. – С. 289-291.
5. Щербич А.Ю. Некоторые вычислительные методы контроля качества геолого-геофизической информации при построении корпоративных банков данных / А.Ю. Щербич // Нефть и газ 2016: сб. тезисов докладов 70-й международной молодежной научной конференции. Т. 1. – М.: Изд-во РГУ нефти и газа (НИУ) имени И.М. Губкина, 2016. – С. 113.

### Публикации в сборниках научных статей

6. Щербич А.Ю. Разработка алгоритма фильтрации случайных шумов на сейсмических трассах / А.Ю. Щербич // Математическое и информационное моделирование: сб. научных трудов. Вып. 9. – Тюмень: Изд-во «Вектор Бук», 2009. – С. 224-232.
7. Щербич А.Ю. Проблемы качества приема и подготовки сейсмических данных в аналоговой и цифровой форме. Передача сейсмички в банки данных / А.Ю. Щербич, В.Н. Кутрунов // Математическое и информационное моделирование: сб. научных трудов. Вып. 10. – Тюмень: Изд-во «Вектор Бук», 2008. – С. 173-182.
8. Щербич А.Ю. Некоторые вычислительные методы контроля качества сейсморазведочной информации при построении корпоративных банков данных / А.Ю. Щербич // Современные технологии нефтегазовой геофизики: материалы докладов конференции 18-19 мая 2016. Отв. редактор С.К. Туренко. – Тюмень: ТИУ, 2016. – С. 100-109.

### Свидетельства о регистрации программ для ЭВМ

9. Щербич А.Ю. Свидетельство об официальной регистрации программы для ЭВМ № 2015614190 «Программа контроля качества метаинформации сейсмических файлов на основе нейронной сети Хемминга «Smart SEG-Y» от 08.04.2015.
10. Щербич А.Ю. Свидетельство об официальной регистрации программы для ЭВМ № 2016610147 «Модуль построения невыпуклых оболочек на множестве отрезков «Concave Hull Builder» от 11.01.2016.

Подписано в печать 12.10.2016. Тираж 120 экз.  
Объем 1,0 уч.-изд. л. Формат 60×84/16. Заказ 870.

---

Издательство Тюменского государственного университета  
625003, г. Тюмень, ул. Семакова, 10.  
Тел./факс (3452) 59-74-81; 59-74-68  
E-mail: [izdatelstvo@utmn.ru](mailto:izdatelstvo@utmn.ru)