КУРСОВАЯ РАБОТА

На тему:

**Реализация компонентов информационной системы архива спутниковых данных**

Владивосток, 2008

# Содержание

Содержание

1. Введение

1.1 Глоссарий

1.2 Описание предметной области

1.3 Неформальная постановка задачи

1.4 Обзор существующих методов решения

2. Требование к окружению

3. Спецификация данных

4. Функциональные требования

5. Проект

5.1 Средства реализации

5.2 Модули и алгоритмы

5.3 Структуры данных

6. Реализация и тестирование

Заключение

Список литературы

# Введение

## 1.1 Глоссарий

* Метаданные – «данные о данных». Краткое структурированное описание, содержащее информацию об определённом ресурсе.
* Система метаданных (информационная система) – программный комплекс, включающий в себя базу метаданных, каталог данных на его основе, сопутствующие интерфейсы доступа и ряд обслуживающих утилит
* Каталог спутниковых данных – система, основанная на метаданных, предоставляющая пользователю информацию о файлах спутниковых данных организации
* База метаданных – база данных, содержащая записи метаданных
* Система инвентаризации (инвентарь) – часть системы метаданных, занимающаяся регистрацией размещения файлов и имеющая набор интерфейсов для доступа, в первую очередь автоматическом режиме
* Утилита инвентаризации – консольное приложение, предназначенное для регистрации файла в системе инвентаризации.
* Продукт – файл исходных данных или конечной продукции, который может быть зарегистрирован в системе.
* Архив – ресурс на файловом сервере, доступный по протоколам SMB (а также FTP/HTTP), содержит файлы продуктов.
* Идентификатор (файла) – строка определённого формата, однозначно идентифицирующая продукт Центра.
* Java Runtime Environment **–** исполнительная среда Java
* ИК канал – Инфракрасный канал

**1.2 Описание предметной области**

В настоящее время спутниковые данные широко используются как источник информации, необходимой для поддержки научно-исследовательской, хозяйствено-экономической, а также военно-политической деятельности человека на Земле. Количество спутников, поставляющих такого рода данные, непрерывно возрастает, также как и количество спутниковых центров, занимающихся приемом и обработкой этой информации.

Одним из таких спутниковых центров является Центр колективного пользования Регионального Спутникового Мониторинга Окружаюшей Среды ДВО РАН. Центр функционирует на базе Лаборатории спутникового мониторинга ИАПУ ДВО РАН [1].

Пользователями, которых интересует продукция спутниковых центров, обычно являются [2]:

* Исследователи природных объектов Земли
* Разработчики методов и средств измерений и исследований
* Природохранные организации

Основными целями пользователей является получение набора значений физических параметров, восстановленных на основе спутниковых данных, а также однозначная идентификация определённых объектов на изображениях. Типовой сценарий работы клиента спутникового центра включает в себя: выбор данных и типа их обработки; заказ обработки и получение её результатов.

Список базовых задач, решаемых спутниковыми центрами, обычно включает в себя [2]:

* Приём данных со спутников
* Хранение архива данных за некоторый промежуток времени
* Информационно-справочный сервис
* Обработка данных различного уровня, включая начальную коррекцию и тематическую обработку
* Доставка данных потребителю (как исходных, так и обработанных)

Вследствие огромного потока разнородной информации возникает множество проблем, связанных с хранением, поиском и получением конкретных данных, необходимых потребителю. В решении основной массы этих задач помогают метаданные. Обычное назначение метаданных – обеспечение пользователя исчерпывающей информацией о спутниковых изображениях архива как в целом, так и по конкретному изображению. Наличие этой информации позволяет пользователю быстро найти данные, пригодные для решения его задач.

Современной тенденцией развития спутниковых центров является их объединение в единые информационные сети. Это обусловлено неодинаковыми возможностями центров по работе с данными того или иного вида и необходимостью использовать данные других организаций. Основными проблемами такого обмена данными являются отличающиеся интересы участников, необходимость согласования стандартов обмена, а также вопросы оплаты и ограничений на применение данных.

Результатом выработки единой стратегии решения перечисленных проблем в Европейском космическом агентстве (ESA) явилась система SSE (Service Support Environment). Прототип системы был отлажен в рамках проекта MASS-ENV, существующего с 2001 года. Разработкой концепции, стандартов и ПО занимается группа компаний под предводительством Spacebel (Бельгия). Проект основан на открытых и общепризнанных технологиях, таких как Web Services и SOAP. Задачей системы является создание среды, объединяющей все службы дистанционного зондирования Европы и других стран.

**1.3 Неформальная постановка задачи**

Существующая в Центре информационная система включает базу метаданных для файлов данных спутников NOAA, веб-интерфейс и утилиту для сбора метаданных из файлов данных. Помимо этого на базе пакета SSE Toolbox осенью 2007 года был создан каталог данных, интегрированный в тестовом режиме в среду SSE. Встаёт задача интеграции наиболее важных ресурсов Центра в данную среду. Помимо этого, развитие системы распределённой обработки спутниковых данных, действующей в Центре, требует развития базы метаданных и программных интерфейсов для доступа к ней в рамках Центра.

На рис. 1 представлена архитектура разрабатываемой информационной системы:

1. Сбор информации о местоположении файла
2. Занесение информациив каталог (внешний URL файла)
3. Извлечение всей потенциально пригодной информации из файла и занесение её БД соответствующего типа
4. Отображение метаданных на каталог в соответствии с требованиями целевой системы
5. Интерфейс в одну из внешних систем

*Каталог данных*

*Метаданные файлов определённого типа*

*Инвентарь*

*Внешняя система*

*Архив*

**1**

**3**

**4**

**2**

**5**

Рис. 1 Архитектура разрабатываемой информационной системы.

На данном этапе наиболее важной является реализация процедур сбора метаданных; а также организация системы инвентаризации, которая упростила бы поиск файлов спутниковых данных, находящихся на различных ресурсах (сетевых хранилищах и CD/DVD носителях).

Целью данной работы является:

* Создание набора классов, реализующих функции генерации метаданых для заданного файла данных спутника MTSAT-1R.
* Реализация системы инвентаризации.

## 1.4 Обзор существующих методов решения

В данный момент в Центре функционирует информационная система для файлов данных, полученных со спутников серии NOAA [3], она позволяет производить поиск и просмотр информации как через веб-интерфейс на сайте Центра, так и через сервер SSE Portal. Реализована утилита извлечения метаданных из файлов данных, полученных с полярно-орбитального спутника FY-1D.

Имеющиеся утилиты извлечения метаданных не могут быть использованы для файлов данных полученных со спутника MTSAT-1R, так как они имеют другой формат.

Новая архитектура системы метаданных потребовала создание подсистемы, отвечающей за справочную информацию о расположении и способе доступа к файлам спутниковых данных. До сих пор эта информация хранилась в базе метаданных наряду с остальными атрибутами. Решение этой задачи в рамках отдельной подсистемы (инвентаря) делает модернизацию и поддержку проще, позволяя компонентам эволюционировать независимо.

**2. Требование к окружению**

Минимальные аппаратные требования включают оперативную память 64 Мб, Процессор с частотой 500 МГц.

Также для работы программы необходима установка Java Runtime Environment версии не ниже 1.5.

Для инвентаризатора также необходима установка сервера PostgreSQL, JDBC драйвер для PostgreSQL.

**3. Спецификация данных**

Процедур сбора метаданных работает с данными, полученными со спутника MTSAT-1R и сохраненных в формате HIRID.

Файл данных состоит из кадров. Каждый кадр представляет собой данные описывавшие этот кадр, и значения полученные радиометрами спутника.

**4. Функциональные требования**

Процедур сбора метаданных требует:

* Методы для извлечения информации, используемой для заполнения полей метаданных
* Методы для генерации уменьшенного изображения

Изображение должно строиться по данным одного из каналов (ИК) и быть кодировано в формат grayscale JPEG и доступно в виде массива байт.

Модули, генерирующие метаданные на основе спутниковых данных должны реализовывать интерфейс MetadataSource. Этот интерфейс описывает произвольный источник метаданных для файлов полярно-орбитальных спутников. Интерфейс требует определения ряда методов для получения необходимой информации. Помимо этих методов, класс, реализующий данный интерфейс, должен иметь конструктор с файлом в качестве параметра (java.io.File).

Для инвентаризации необходимо:

* База данных инвентаризации.
* Утилита, заполняющая базу данных инвентаря данными о расположении файла.
* Утилита для извлечения этих данных из инвентаря.

**5. Проект**

## 

## 5.1 Средства реализации

Так как в связи с спецификой лаборатории необходима работа под любой платформой и существующая часть системы написана на языке Java в качестве языка программирования был выбран язык Java. Сервером базы данных был выбран PostgreSQL так как он уже функционирует в лаборатории.

Инструментарий разработки:

* Sun JDK 1.6 – компилятор + утилиты и библиотеки
* IntelliJ IDEA 7 – интегрированная среда разработки
* PostgreSQL 8.2 – сервер базы данных
* pgAdmin 3 – утилита для администрирования СУБД PostgreSQL

## 5.2 Модули и алгоритмы

Проект состоит из двух классов MTSATMetadataSource и MTSAT.

MTSATMetadataSource – реализует интерфейс MetadataSource. Внутри конструктора создаются экземпляр. Реализованы методы возвращающие необходимые поля этого экземпляра.

MTSAT – отвечает за получения необходимых полей метаданных на основе информации из заголовка файла, в том числе координаты границ изображения, даты и времени приёма, виток спутника, числа строк и присутствующие каналы.

MetadataSource

MTSATMetadataSource

MTSAT

Таблица 1. Основные методы класса MTSATMetadataSource

|  |  |
| --- | --- |
| **Метод** | **Описание** |
| T0MetadataSource | Конструктор класса, на входе имеет файл |
| getFormat getSatellite getSensors getTime getLu getRu getRd getLd getLength getTurn getQuality getBlock0 getPreview | Методы для чтения значений полей. Вызываются извне и необходимы для функционирования класса в рамках системы |

Таблица 2. Основные методы класса MTSAT

|  |  |
| --- | --- |
| **Метод** | **Описание** |
| MTSAT | Конструктор класса, на входе имеет файл |
| setSatelliteId setChannels setSatelliteTurn setQualitysetLatRu setLonLu setLength setLonRu setLatRd setLonRd setLatLd setLonLd | Методы для установки значений соответствующих полей. Вызываются изнутри |
| getId getFile getSatelliteId getSatelliteTurn getChannels getLength getQuality getLatLu getLonLu getLatRu getLonRu getLatRd getLonRd getLatLd getLonLd | Методы для чтения значений полей. Вызываются извне и необходимы для функционирования класса в рамках системы |

Для инвентаризатора были реализованы две утилиты INV и INF.

INV- получает на вход полное имя файла, архив в котором он хранится и тип спутника. Заполняет базу данных инвентаря данными о файле и выдает идентификатор файла.

INF-получает на вход идентификатор файла. Выдает информацию о файле.

Таблица 3. Основные методы класса MTSATMetadataSource

|  |  |
| --- | --- |
| **Метод** | **Описание** |
| getID | Вход дата и время. Выход идентификатор |
| getDate | Возвращает дату. |
| Get Time | Возвращает время. |

## 5.3 Структуры данных

База данных состоит из 4 таблиц file, folder, archive, file2folder.

file – Содержит информацию о файле

folder – Содержит информацию о размещении

archive – Содержит информацию об архиве

file2folder – связывает файл с его размешением

ID

Name

invDate

createDate

modifDate

size

fileID

folderID

ID

archiveID

path

ID

Name

label

file

file2folder

folder

archive

Таблица 4. таблица file

|  |  |  |
| --- | --- | --- |
| Название поля | Тип | Описание |
| Id | Text | Идентификатор файла |
| Name | Text | Имя файла |
| invDate | Timestamp | Дата-время инвентаризации файла |
| createDate | Timestamp | Дата-время создания файла |
| modifDate | Timestsmp | Дата-время изменения файла |
| Size | Int8 | Размер файла |

Таблица 5. таблица folder

|  |  |  |
| --- | --- | --- |
| Название поля | Тип | Описание |
| Id | Int4 | Идентификатор папки |
| Path | Text | Путь к папке |
| archivID | Int4 | Идентификатор архива в котором находиться папка |

Таблица 6. таблица archive

|  |  |  |
| --- | --- | --- |
| Название поля | Тип | Описание |
| Id | Text | Идентификатор архива |
| Name | Text | Имя архива |
| Label | Text | Описание |

Таблица 7. таблица file2folder

|  |  |  |
| --- | --- | --- |
| Название поля | Тип | Описание |
| fileId | Text | Идентификатор файла |
| folderId | Int4 | Идентификатор папки в которой лежит файл |

**6. Реализация и тестирование**

Исходный код классов занимает 20 кб (примерно 500 строк кода на языке Java и примерно 50 строк на SQL). Код расположен в пяти файлах MTSAT.java, MTSATMetadataSource.java, INV.java, INF.java inventor.sql .

Механизм вызова модуля, соответствующего типу файла, в настоящее время не реализован, поэтому для тестирования сбора метаданных был использован следующий упрощенный код:

Class aClass = Class.forName("em.bv.storage.MTSATMetadataSource");

Constructor constructor = aClass.getConstructor(File.class);

MetadataSource source = (MetadataSource) constructor.newInstance(newFile("g5200803010133.gms"));

System.out.println("time="+ source.getTime());

System.out.println("length=" + source.getLength());

System.out.println("quality=" + source.getQuality());

System.out.println("source.getLu() = " + source.getLu().toTrimString());

System.out.println("source.getRu() = " + source.getRu().toTrimString());

System.out.println("source.getRd() = " + source.getRd().toTrimString());

System.out.println("source.getLd() = " + source.getLd().toTrimString());

System.out.println("satelliteTurn=" + source.getTurn());

System.out.println("preview length=" + source.getPreview().length);

System.out.println("blk0 length=" + source.getBlock0().length);

При тестировании системы инвентаризации были использованы имеющиеся файлы, заполнена база данных инвентаризации. Получена информация о каждом файле, хранящемся в этой базе.

В ходе тестирования ошибок обнаружено не было.

# Заключение

Таким образом, в процессе курсовой работы:

* Был изучен формат файла данных со спутника MTSAT-1R
* Был реализован алгоритм получения необходимых полей с нужными данными
* Была разработана структура базы данных инвентаря
* Реализована утилита внесения данных в инвентарь
* Реализована утилита получения данных из инвентаря

# Список литературы

1. Центры коллективного пользования Российской Академии Наук // М.: Наука, 2004, 192 с.
2. Ефремов В.Ю., Лупян E.A., Мазуров A.A., Прошин А.А., Флитман Е.В. Технология построения автоматизированных систем хранения спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей. Москва: Полиграф сервис, 2004 с. 437-443
3. Каталог метаданных для файлов NOAA // Центр регионального спутникового мониторинга окружающей среды ДВО РАН // http://www.satellite.dvo.ru