

# Инструменты анализа спутниковых наблюдений природных пожаров на различных типах территорий

Кобец Д.А., Балашов И.В.

Распределенные архивы, созданные и поддерживаемые ИКИ РАН [3], содержат долгосрочные временные ряды данных дистанционного зондирования как зарубежных («Aqua», «Landsat», «Noaa», «Sentinel», «Terra», ... ) так и отечественных («Канопус», «Метеор», «Ресурс») спутниковых систем. Накопленные многолетние массивы данных позволяют решать широкий спектр научных задач, в частности создавать ежегодно обновляемые карты растительного покрова [1] и преобладающих древесных пород [3] Российской Федерации, осуществлять мониторинг лесных пожаров [6].

Информационные продукты, созданные на основе спутниковых данных, в том или ином виде, представляют собой географически привязанную информацию, хранящуюся в разнообразных системах, осуществляющих хранение данных как в файлах различных форматов и типов, так и во множестве разнородных БД.

Таким образом, одними из основных типов интерфейсов работы с такой информацией являются WEB ГИС интерфейсы. Подобные интерфейсы предоставляют пользователю возможность работы со спутниковыми данными и продуктами на их основе, имея минимум необходимых для этого ресурсов (достаточно

среднестатистического компьютера с возможностью выхода в интернет), и предназначены для анализа пространственного расположения различных объектов и явлений в определенный момент времени, но при этом не предоставляют расширенного функционала позволяющего оценить эти пространственные распределения количественно. Кроме того, в последнее время встают задачи проведение не только анализа одномоментных характеристик объектов и явлений, но и их временных рядов. Так как распределенные системы хранения ориентированы на ведение архивов картографической информации, а так же временных и пространственных характеристик динамики наблюдаемых объектов и явления, то они располагают всеми необходимыми данными для проведения подобного анализа.

Следует отметить, что в последнее десятилетие достаточно быстро развивались различные технологии связанные с созданием подобных инструментов для проведения анализа различной информации. Одними из наиболее быстро развивающихся в этом направлении технологий создания подобных инструментов являются BI-технологии (Business intelligence) [5].



Одной из основных особенностей систем BI-технологий является то, что они предоставляют в распоряжение пользователя обширный инструментарий [2], позволяющий ему прямо из WEB-интерфейса запросить необходимые для анализа данные, и просматривает полученные результаты через динамически формируемые интерактивные отчеты. Интерактивные отчеты предоставляют пользователю возможность самому решать, какие именно данные требуются для решения стоящих перед ним задач анализа и как эти данные необходимо организовать. Таким образом, имея в своем распоряжении сравнительно небольшое количество интерактивных отчетов, пользователь способен решать с их помощью широкий спектр аналитических задач (перестраивая их под свои нужды).

Исходные данные, анализ которых необходимо произвести, в общем случае могут располагаться в различных системах, осуществляющих хранение данных как в файлах различных форматов и типов, так и во множестве разнородных БД. Чтобы сделать данные из этих источников доступными для анализа, их необходимо консолидировать.

Первым этапом консолидации является применение к данным методов позволяющих извлечь их из источников и перенести в специализированную систему хранения.

Извлеченные данные могут не иметь приемлемого уровня качества, поэтому их не достаточно просто разместить в структурированном источнике, необходимо также произвести комплекс методов и процедур направленных на очистку данных от факторов не позволяющих производить их корректный анализ, а так же на обогащение их дополнительной информацией (необходимой для решения определенных аналитических задач).

Консолидированные данные, размещенные в реляционных БД и описывающие

какие-либо объекты или процессы, зачастую содержат множество показателей, свойств и атрибутов, которые, с точки зрения анализа, можно охарактеризовать как информационные измерения. Располагая в реляционных системах хранения, такие данные (будучи нормализованными) хранятся в большом количестве таблиц, связанных некоторым набором отношений, что снижает скорость выполнения аналитических запросов, а так же делает такие данные сложными для визуального анализа и осмысления.

Для решения связанных с реляционными БД проблем, в системах BI-аналитики используются хранилища основанные на многомерной модели данных опирающихся на концепцию OLAP-кубов (англ. On-Line Analytical Processing, оперативная аналитическая обработка).

Данные формирующие OLAP куб делятся на два вида:

- «Измерения» – характеристики, описывающие исследуемый объект или процесс, принимающие значения из ограниченного набора (т.е. являющиеся дискретными).

- «Факты» – некоторые показатели (численные значения), которые соответствуют исследуемому объекту или процессу, полученные при фиксированных значениях (диапазонах значений) измерений.

OLAP-куб (в общем случае) формируется под конкретную интерактивную отчетную форму. Интерактивная отчетность является гибким инструментом, настраиваемым под решение широкого круга задач. Работу любой отчетной формы, обеспечивает её прототип, представляющий собой XML-файл разметки, содержащий в себе описание всех необходимых для её функционирования параметров (источников данных, методов (и их параметров) осуществляющих извлечение данных из источников, правила трансформации извлеченных данных (сортировка, взаимосвязь с другими наборами данных, присвоение дополнительных параметров (например цвета) и т.п.), правила формирования многомерных кубов из полученных данных, функции свертки и декомпозиции фактов в иерархии измерений, правила формирования дополнительных фактов, параметры визуализации данных).

Так как зачастую сервера источников данных, специализированные системы хранения консолидированных данных и системы, интерпретирующие прототипы интерактивных отчетов, преобразующие данные к многомерному виду, физически размещаются на различных серверах, это позволяет создавать распределенные системы анализа данных на основе BI-технологий.

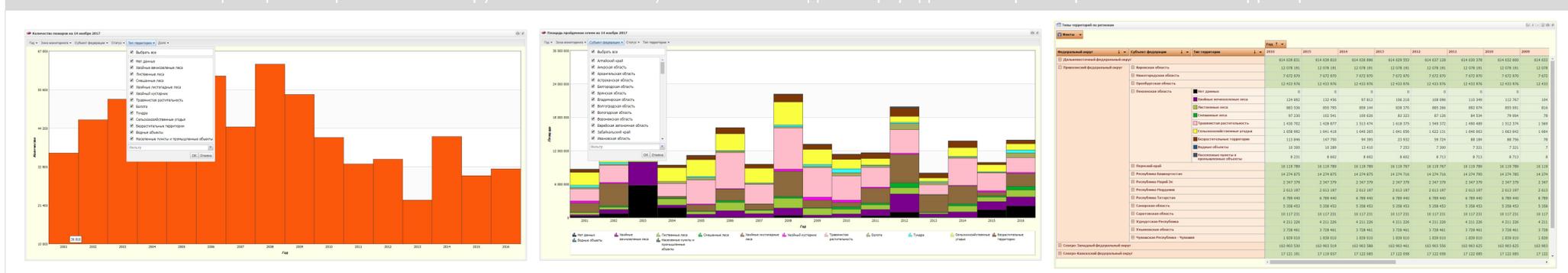
Возможность определять диапазон данных для формирования интерактивной отчетной формы и возможность работать с этими данными, позволяет создавать формы не под конкретные задачи, а под целые классы задач, данные для которых могут извлекаться из различных источников, находящихся в территориально разделенных друг от друга центрах.

Так как архивы наблюдаемых пожаров достаточно хорошо структурированы, то используя данные полученные в результате классификации различных видов растительного покрова (за предыдущий сезон) в совокупности с информацией о пространственных характеристиках пожаров, возможно производить количественный анализ различных типов растительного покрова пострадавших от огня, а так же оценивать временную динамику происходящих на них пожаров.

Интерактивные отчетные формы, по средствам которых производится такой анализ, предоставляют пользователю возможность: организовать доступ к данным через WEB-интерфейс; формировать отчет под необходимый пользователю временной диапазон исследуемых характеристик; фильтровать данные по субъектам федерации, годам, зонам мониторинга и типам территорий; выгружать результат в различные форматы («.png», «.pdf», «.xlsx», «.docx», «.html»).



## Примеры интерактивных инструментов анализа спутниковых наблюдений природных пожаров на различных типах территорий



Работа выполнена при поддержке Минобрнауки России (инструменты оценки характеристик лесных ресурсов), контракт № 14.607.21.0122).



1



2



3



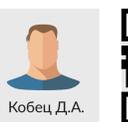
4



5



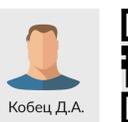
6



8



9



Кобец Д.А.



ИКИ  
ИНСТИТУТ  
КОСМИЧЕСКИХ  
ИССЛЕДОВАНИЙ  
РАН