

Использование BI-технологий для создания инструментов в анализе данных спутникового мониторинга

**Д.А. Кобец, И.В. Балашов, И.Д. Данилов,
Е.А. Лупян, И.Г. Сычугов, В.А. Толпин**

*Институт космических исследований РАН
Москва, 117997, Россия
E-mail: kobets@d902.iki.rssi.ru*

В работе рассматривается опыт применения технологий Business intelligence (BI) для построения web-интерфейсов, обеспечивающих возможность проведения анализа информации, получаемой на основе спутниковых данных в системах дистанционного мониторинга. Рассмотрены основные особенности и возможности BI-технологий. Приведен анализ различных типов информации, получаемой на основе данных дистанционного зондирования в системах мониторинга, и основные особенности организации работы с ними. На основе проведенного анализа предложена архитектура схемы внедрения в информационные системы дистанционного мониторинга блоков, построенных на основе BI-технологий, обеспечивающих работу с постоянно обновляющейся информацией. Рассмотрены особенности построения этих блоков и приведено описание базовых схем и программного обеспечения их реализации в различных системах дистанционного мониторинга, которые были разработаны в Институте космических исследований РАН. Эти разработки позволяют организовывать в потоковом режиме процессы преобразования поступающей в системы дистанционного мониторинга информации в вид (OLAP-куб), необходимый для эффективной работы с данными при проведении их анализа, в том числе осуществления интеграции информации при построении всевозможных табличных и графических отчетов. В работе также обсуждаются вопросы опыта использования предложенной архитектуры и программного обеспечения для ее реализации, и построения на основе BI-технологий web-интерфейсов в различных системах мониторинга. Данные вопросы обсуждаются на примере использования разработанных подходов в Информационной системе дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ-Рослесхоз) и спутниковом сервисе BEGA-Science, созданном в ИКИ РАН для решения задач дистанционного мониторинга возобновляемых биологических ресурсов.

Ключевые слова: BI-технологии, OLAP, анализ спутниковых данных, интерактивная отчетность, отчетные формы, системы дистанционного мониторинга, дистанционное зондирование

Введение

В настоящее время данные дистанционного зондирования Земли (ДЗЗ) все шире используются для количественной информации о различных объектах и территориях. Современные технологии дистанционного зондирования и построенные на их основе системы дистанционного мониторинга позволяют получать и накапливать как пространственные распределения, так и временные ряды характеристик различных наблюдаемых объектов и явлений. Примерами такой информации являются, например, ряды ежегодно обновляющихся карт растительного покрова, пахотных земель, земель, занятых различными видами посевов, базы данных наблюдений природных пожаров и т.д.

При этом во многих случаях данная информация может быть достаточно хорошо структурированной и на ее основе можно анализировать как пространственную, так и временную динамику различных объектов и явлений. Для проведения такого анализа уже нужны инструменты, не просто обеспечивающие работу с пространственной информацией, но и позволяющие анализировать различные характеристики, полученные на основе обработанной спутниковой информации, например, распределения тех или иных объектов по территориям и административным единицам или временные ряды различных характеристик. Для этого необходимы методики, которые позволяют представить информацию в удобном

для анализа виде, например, в графической, табличной или пространственной форме, а также осуществлять удобный выбор различных параметров для проведения ее 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 являются BI-технологии (Business intelligence) (Moss et al., 2003), базирующиеся на построении хранилищ структурированной информации на основе OLAP (online analytical processing) (Палкин и др., 2013). Эти технологии сегодня позволяют строить достаточно сложные инструменты для онлайн анализа больших объемов данных.

В настоящей статье показано, что предлагаемые в BI-технологии подходы могут достаточно эффективно использоваться при построении инструментов для работы с информацией, получаемой на основе спутниковых данных в системах дистанционного мониторинга. В работе представлен подход, позволяющий создавать и встраивать в такие системы созданные на основе BI-технологий блоки, обеспечивающие возможность анализа получаемой в системах информации. Основной особенностью данного подхода является автоматическое, потоковое структурирование информации, получаемой на основе спутниковых данных, и преобразование ее в вид, позволяющий использовать ее в онлайн интерфейсах, обеспечивающих представление и анализ. В работе обсуждаются основные особенности BI-технологий, которые могут быть эффективно использованы в системах дистанционного мониторинга, разработанная архитектура построения блоков анализа данных и особенности реализации ее блоков. Также описаны возможности основных типов интерфейсов, которые могут создаваться для анализа результатов обработки спутниковой информации, и приведены примеры использования разработанных подходов и базового программного обеспечения для создания блоков аналитической отчетности в различных системах дистанционного мониторинга.

Введение в BI технологий

BI-технологии (Business Intelligence) (Moss et al., 2003) позволяют на основе единого механизма создавать широкий спектр интерактивных инструментов для анализа различной информации (отчетов), представляемой в виде таблиц, графиков и карт. Одной из основных особенностей таких технологий является то, что они дают возможность строить динамичные отчеты. Это позволяет пользователю самостоятельно решать, какие анализируемые данные должны быть ему представлены, а также определять их вид и состав.

В то же время следует отметить, что в различных системах хранение данных может быть организовано таким образом, чтобы оптимизировать их получение и хранение. Оптимизация информации, обеспечивающая интеграцию для проведения различных уровней обработки и анализа, не всегда позволяет организовать эффективные запросы к данным. Поэтому для работы интерфейсов, обеспечивающих возможность быстрого и удобного анализа данных, в том числе их интеграции, во многих случаях необходимо преобразование данных в вид, тре-

буемый для проведения эффективного выбора информации, необходимого для решения характерных задач, возникающих при анализе. В рамках подходов BI для решения данной проблемы используется специальная технология OLAP (OnLine Analytical Processing) (Палкин и др., 2013), заключающаяся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу. OLAP-структура, созданная из рабочих данных, называемая OLAP-куб (рис. 1, где $t_0 - t_n$ – временные срезы массива векторов объектов $F_0 - F_n$, содержащих параметры $P_0 - P_n$), представляющий из себя многомерный массив, созданный из соединенных таблиц базы данных (БД). Особенностью OLAP-структуры является возможность произвести выборку данных из любого среза куба (пунктирные линии, рис. 1). Пожалуй, наиболее важными особенностями современных технологий BI является наличие инструментария, позволяющего достаточно легко формировать OLAP-кубы и различные виды web-интерфейсов для работы с ними.

На сегодняшний день существует множество программного обеспечения для реализации BI-технологий:

- SAP BI/BO – комплекс аналитических приложений (<http://go.sap.com/solution/platform-technology/business-intelligence.html>);
- IBM Cognos BI – комплексное решение для создания информационно-аналитической системы в масштабах всей компании (<http://www-03.ibm.com/software/products/ru/business-intelligence>);
- Oracle Business Intelligence – решение, обеспечивающее доступ ко всем информационным источникам компании (<http://www.oracle.com/us/solutions/business-analytics/business-intelligence/overview/index.html>);
- Contour Components – данное решение было использовано при реализации описанных в статье блоков аналитики (<http://www.contourcomponents.com/ru.htm>).

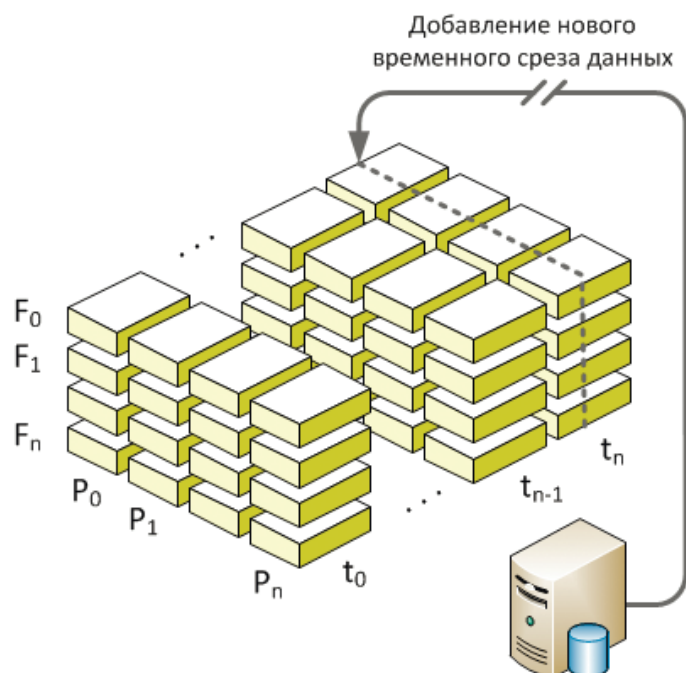


Рис. 1. Схема организации OLAP-кубов

Особенности построения блока VI-аналитики для проведения анализа результатов обработки данных спутникового мониторинга

Наиболее распространенной задачей дистанционных наблюдений является получение информации либо об отдельных объектах, либо о протяженных территориях и явлениях. При этом дистанционные наблюдения позволяют получать как временные, так и пространственные характеристики динамики наблюдаемых объектов и явления. В частности, получаемая на основе дистанционных наблюдений информация позволяет решать задачи:

- анализ динамики площадных характеристик различных объектов (анализ карт) обычно в разрезе территориального деления;
- анализ динамики не площадных характеристик различных объектов.

Следует отметить, что при осуществлении мониторинга различных объектов и явлений системы хранения получаемых данных обычно создаются в виде, оптимальном для ее оперативного обновления и отображения. Создаваемые архивы могут быть ориентированы либо на ведение БД характеристик наблюдаемых объектов, либо на хранение различной картографической информации. В основном такие системы хранения не являются оптимальными для использования в системах аналитики, построенных на основе OLAP. Поэтому для обеспечения эффективной работы блоков VI-аналитики результатами обработки данных дистанционного мониторинга обычно необходимо формирование дополнительной специализированной системы хранения, оптимизированной для быстрого выбора данных и проведения их анализа. При этом следует также учитывать возможности имеющихся в настоящее время специализированных программных решений, реализующих VI-технологии, которые позволяют создавать достаточно удобные и эффективные WEB-интерфейсы для проведения анализа. Фактически, для эффективного использования VI-технологий при создании инструментов анализа данных дистанционного мониторинга, информацию необходимо преобразовывать в вид, удобный для работы систем, обеспечивающих построение WEB-интерфейсов.

Характерная архитектура интеграции блоков VI-аналитики в системы дистанционного мониторинга

Предлагаемая архитектура интеграции блоков VI-аналитики для обеспечения работы с данными дистанционного мониторинга (рис. 2) представлена следующими основными элементами.

1. Блок организации автоматизированного приема и сбора спутниковых данных, который обеспечивает либо непосредственный прием со спутников, либо получение данных различных специализированных центров приема.
2. Блок первичной обработки спутниковых данных – обеспечивает временную и географическую привязку, калибровку, коррекцию на условия освещенности, коррекцию влияния атмосферы, трансляцию данных в определенную географическую проекцию и т.д.

3. Блок архивации данных, обеспечивающий ведение оперативных и долговременных архивов спутниковых данных и результатов их обработки, оптимизированных под организацию процессов сбора обработки и представления конкретной информации.
4. Блок преобразования информации в вид, необходимый для работы BI-аналитики. Осуществляет в автоматическом режиме перерасчет архивов результатов обработки спутниковых данных в OLAP-куб.
5. OLAP-куб – подготовленная для анализа агрегированная информация, структурированная по многомерному принципу, рассчитываемая на основе характеристик отслеживаемых объектов и территорий, хранящихся в долгосрочных архивах спутниковых данных.
6. Блок программного обеспечения BI-аналитики – осуществляет взаимодействие пользователя со срезами данных OLAP-куба по средствам WEB-приложения.

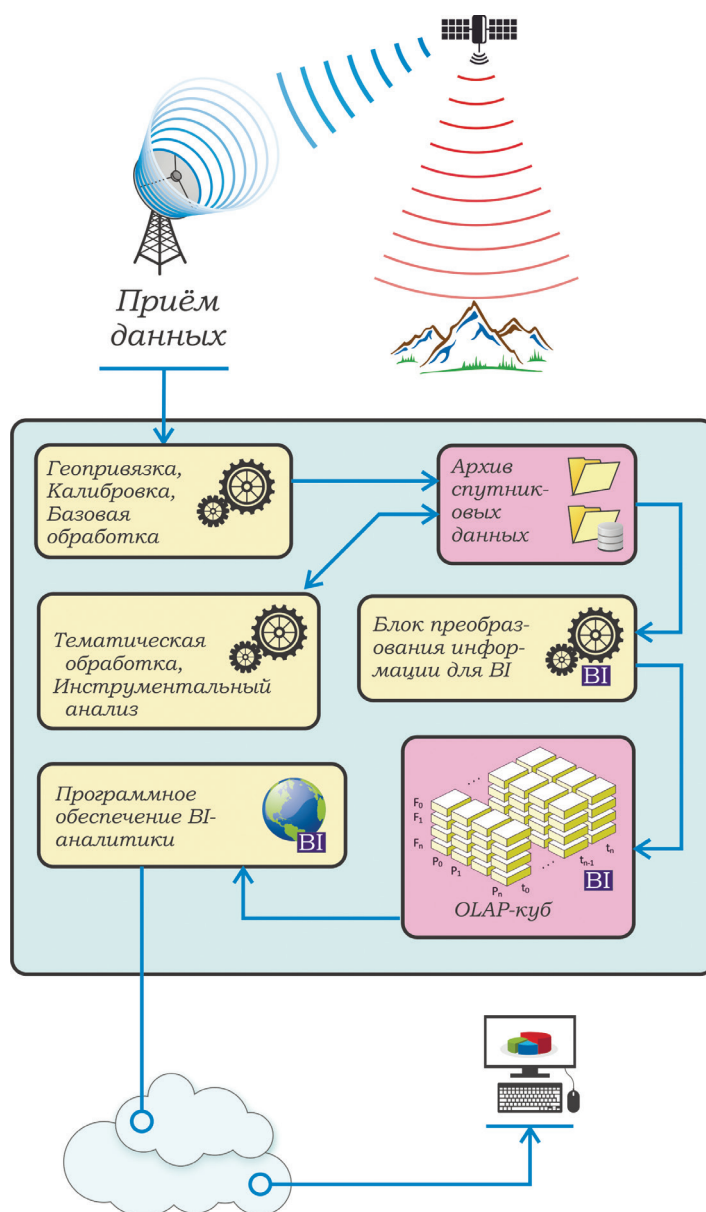
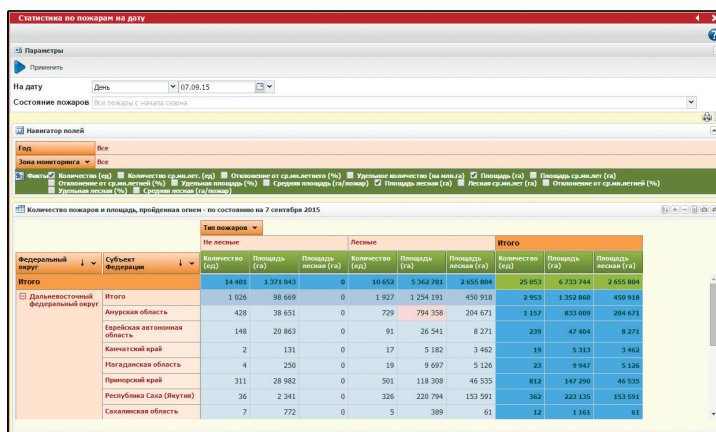
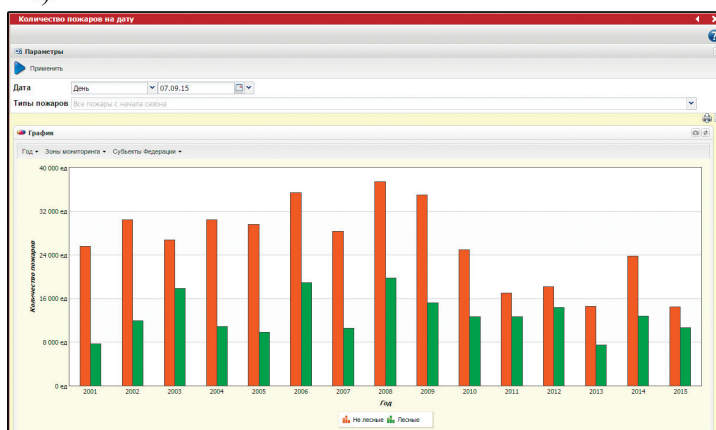


Рис. 2. Характерная архитектура построения блоков BI-аналитики в системах дистанционного мониторинга

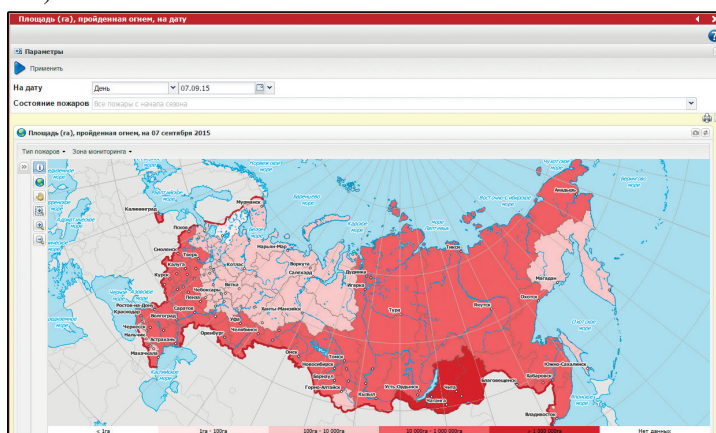
В системах дистанционного мониторинга, поддерживаемых Институтом космических исследований Российской академии наук (ИКИ РАН), предложенная архитектура реализована с использованием технологий и базового программного обеспечения сбора, обработки и распространения спутниковых данных, разработанных в ИКИ РАН (Лупян, Мазуров и др., 2011; Толпин и др., 2011), и специализированного программного обеспечения Contour BI Portal (<http://www.contourcomponents.com/page/ru.101.htm>), позволяющего как организовывать ведение OLAP-куба, так и формировать различные виды интерактивных, аналитических web-интерфейсов для проведения анализа таблиц, графиков и карт (рис. 3).



а).



б).



в).

Рис. 3. Виды интерактивных отчетов: таблицы (а), графики (б), географические карты (в)

Примеры применения разработанных подходов и базового ПО в различных системах мониторинга

Реализованная схема построения блоков VI-аналитики позволила начать создание специализированных интерфейсов для анализа результатов обработки спутниковых данных в различных системах дистанционного мониторинга. Ниже мы приведем несколько примеров таких реализаций.

Информационная система дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ-Рослесхоз) (<http://www.nffc.aviales.ru>) предназначена для комплексного информационного обеспечения работ по мониторингу пожарной опасности, контролю распространения и тушения лесных пожаров, а также оценке их последствий. С основными возможностями системы можно, в частности, познакомиться в работах (Барталев и др., 2008, 2010). Созданный в рамках ИСДМ-Рослесхоз блок аналитики позволяет пользователям осуществлять динамический анализ оперативной и исторической информации о природных пожарах, наблюдающихся на территории Российской Федерации.

Следует отметить, что архивы регистрируемых пожаров имеют достаточно сложную структуру. Для того, чтобы реализовать удобные интерфейсы, обеспечивающие высокую скорость работы с данными, был организован перерасчет данных из архивов регистрируемых пожаров в специализированную БД (OLAP-куб), в которой информация организована таким образом, чтобы можно было осуществлять ее быструю интеграцию по времени и различным видам территорий (типам преобладающей растительности, административному делению, зонам мониторинга и т.д.) Это позволило обеспечить возможность достаточно быстрой работы, в частности, обеспечить одновременный анализ как оперативных, так и исторических данных (рис. 4).

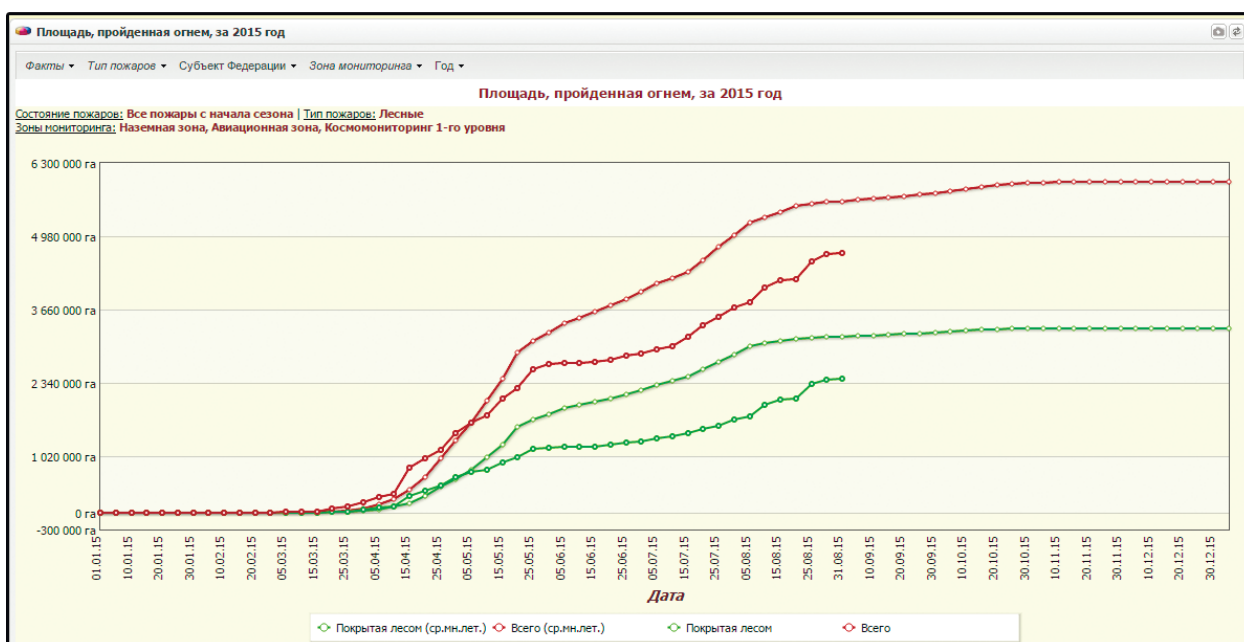


Рис. 4. Интерфейс анализа лесопожарной обстановки, где одновременно отображаются оперативные данные 2015 г. и средние значения динамически получаемые за период 2000–2015 гг.

При этом следует иметь ввиду, что для построения подобных отчетов система оперирует информацией о всех пожарах на территории России, наблюдавшихся по спутниковым данным с начала 2000 года по настоящее время, т.е. осуществляет анализ информации более чем о 670 000 пожарах.

Сервис анализа данных спутниковых наблюдений для оценки и мониторинга возобновляемых биологических ресурсов BEGA-Science (<http://sci-vega.ru/>) (Лупян и др., 2011), также имеет в своем составе блок BI аналитики. Данный блок, в частности, позволяет пользователям проводить динамический анализ ежегодно обновляющейся информации о растительном покрове на территории Российской Федерации (Барталев, Егоров и др., 2011). Для этого в сервисе реализована схема формирования OLAP-куба на основе информации, хранящейся в системе в виде растровых карт. Карты фрагментируются на участки, на которых считаются характеристики типов растительности (рис. 5), и на основе рассчитанных характеристик происходит формирование OLAP-куба. Такая схема позволяет проводить динамический анализ информации на всей территории страны. Отметим, что для решения задач, связанных с анализом ситуации в различных субъектах РФ, интеграция информации при формировании OLAP-куба производится в системе не по равномерной сетке, а в границах субъектов РФ. С примерами интерфейсов, работу которых обеспечивает данный блок, можно ознакомиться по адресу <http://sci-vega.ru/documents/toproxy/bi/vegas/>.

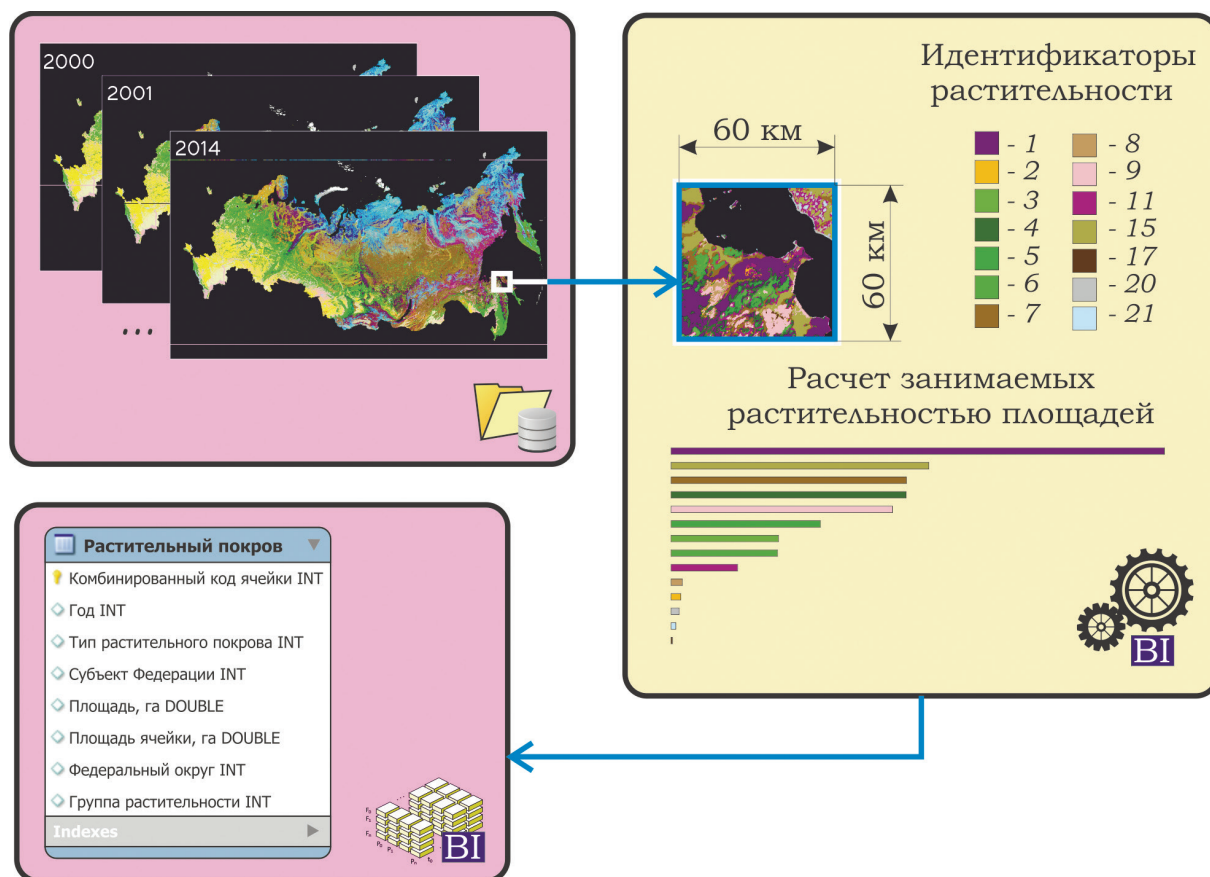


Рис. 5. Схема формирования OLAP-куба по данным карт растительного покрова, полученных на основе спутниковых данных

Для зарегистрированных пользователей в системе доступны различные типы интерфейсов, ориентированные на проведение анализа различных характеристик состояния растительного покрова. Отметим, что система интерфейсов, построенных на основе VI-технологий в сервисе ВЕГА-Science и ориентированных на анализ состояния и динамики растительного покрова, реализована при поддержке Российского научного фонда (проект 14-17-00389).

Заключение

Таким образом, мы видим, что при работе с различной информацией, получаемой в современных системах дистанционного мониторинга, для построения web-интерфейсов, обеспечивающих анализ спутниковых данных дистанционного зондирования, могут достаточно эффективно использоваться современные VI-технологии. Это позволяет реализовать динамические интерфейсы для анализа информации, необходимой при решении научных и прикладных задач. При этом создание и интеграция в системы дистанционного мониторинга специализированных блоков, обеспечивающих функционирование интерфейсов и анализ оперативно поступающих данных, возможны на основе предложенных в работе подходов и созданного в ИКИ РАН базового программного обеспечения. Удобство, устойчивость и эффективность была проверена авторами на примере реализации подобных блоков в различных информационных системах.

Работа выполнялась при поддержке Российской академии наук (тема «Мониторинг», госрегистрация № 01.20.0.2.00164). Разработка использованных в работе карт растительного покрова Российской Федерации, а также создание на основе VI-технологий в среде сервиса ВЕГА-Science интерфейса доступа к полученной с их использованием информации выполнено в ИКИ РАН за счет гранта Российского научного фонда (проект №14-17-00389).

Литература

1. Барталев С.А., Егоров В.А., Ершов Д.В., Исаев А.С., Лупян Е.А., Плотников Д.Е., Уваров И.А. Спутниковое картографирование растительного покрова России по данным спектрорадиометра MODIS // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 4. С. 285–302.
2. Барталев С.А., Ершов Д.В., Коровин Г.Н., Котельников Р.В., Лупян Е.А., Щетинский В.Е. Информационная система дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства РФ (состояние и перспективы развития) // Современные проблемы дистанционного зондирования Земли из космоса. 2008. Выпуск 5. Т. 2. С. 419–429.
3. Барталев С.А., Ершов Д.В., Коровин Г.Н., Котельников Р.В., Лупян Е.А., Щетинский В.Е. Основные возможности и структура информационной системы дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ Рослесхоз) // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7. № 2. С. 97–105.
4. Ефремов В.Ю., Балашов И.В., Котельников Р.В., Лупян Е.А., Мазуров А.А., Прошин А.А., Толпин В.А., Уваров И.А., Флитман Е.В. Объединенный картографический интерфейс для работы с данными ИСДМ-Рослесхоз // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 3. С. 129–139.
5. Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В., Крашенинникова Ю.С. Технологии построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 26–43.

6. Лупян Е.А., Савин И.Ю., Барталев С.А., Толпин В.А., Балашов И.В., Плотников Д.Е. Спутниковый сервис мониторинга состояния растительности («Вега») // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С.190-198.
7. Палкин Н., Орешков В. Бизнес-аналитика. От данных к знаниям // OLAP-анализ. С.-Петербург: Питер, 2013. С. 184–192.
8. Толпин В.А., Балашов И.В., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Уваров И.А., Флитман Е.В. Создание интерфейсов для работы с данными современных систем дистанционного мониторинга (система GEOSMIS) // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 3. С. 93–108.
9. Moss L.T., Atre S.. Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Addison-Wesley, 2003. 576 p.

Using BI technologies to create analysis tools for satellite remote sensing data

**D.A. Kobets, I.V. Balashov, I.D. Danilov,
E.A. Loupian, I.G. Sychugov, V.A. Tolpin**

*Space Research Institute RAS, Moscow 117997, Russia
E-mail: kobets@d902.iki.rssi.ru*

The paper describes an experience of using Business Intelligence (BI) technologies to build web-interfaces providing analysis of satellite data obtained by remote monitoring systems. Main features and opportunities of BI-technologies are considered. Various types of information derived from remote sensing data in monitoring systems and main information management issues are analyzed. An architecture is suggested to introduce BI-technology blocks responsible for continuous information flow management into remote monitoring systems. We discuss specific block development features, basic layouts and implementation software employed in different remote monitoring systems elaborated by Space Research Institute RAS (IKI RAS). These solutions enable organizing stream processes of satellite data conversion into OLAP structures best suitable for data analysis, including information integration when compiling different graphic reports (tables, maps and charts). Our experience of using the suggested architecture and software and developing web-interfaces based on BI-technologies is demonstrated by the examples of Forest Fires Satellite Remote Monitoring System of Federal Forestry Agency and satellite vegetation monitoring service VEGA-Science developed by IKI RAS.

Keywords: BI-technologies, OLAP, satellite remote sensing data analysis, interactive reporting, reporting forms, satellite remote monitoring systems, remote sensing

References

1. Bartalev S.A., Ershov D.V., Ershov D.V., Isaev A. S., Loupian E.A., Plotnikov D.E., Uvarov I.A., Sputnikovoe kartografirovaniye rastitel'nogo pokrova Rossii po dannym spektrometri (MODIS Mapping of Russia's vegetation cover using MODIS satellite spectroradiometer data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 4, pp. 285–302.
2. Bartalev S.A., Ershov D.V., Korovin G. N., Kotelnikov R.V., Loupian E.A., Shetinskiy V. E., Informatsionnaya sistema distantsionnogo monitoringa lesnykh pozharov Federal'nogo agentstva lesnogo khozyaistva RF (sostoyaniye i perspektivy razvitiya) (Forest fires remote monitoring systems by Federal forestry Agency of The Russian Federation (the state and development prospects)), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2008, Issue 5, Vol. 2, pp. 419–429.
3. Bartalev S.A., Ershov D.V., Korovin G.N., Kotelnikov R.V., Loupian E.A., Tshetinskii V.E., Osnovnye vozmozhnosti i struktura informatsionnoi sistemy distantsionnogo monitoringa lesnykh pozharov Federal'nogo agentstva lesnogo khozyaistva (ISDM Rosleskhoz) (The main functionalities and structure of the Forest Fire Satellite Monitoring Information System of Russian Federal Forestry Agency (SMIS-Rosleskhoz)), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2010, Vol. 7, No. 2, pp. 97–105.
4. Efremov V. Yu., Balashov I.V., Kotelnikov R. V., Loupian E.A., Mazurov A. A., Proshin A. A., Tolpin V. A., Uvarov I.A., Flitman E. V., Ob"edinennyyi kartograficheskii interfeis dlya raboty s dannymi ISDM-Rosleskhoz (Integrated mapping interface for operations with ISDM-Rosleskhoz data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 129–139.
5. Loupian E.A., Mazurov A. A., Nazirov R. R., Proshin A. A., Flitman E. V., Krashenninnikova Yu.S., Tekhnologii postroeniya informatsionnykh sistem distantsionnogo monitoringa (Technologies for building remote monitoring

- information systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 26–43.
6. Loupian E.A., Savin I. U., Bartalev S.A., Tolpin V. A., Balashov I.V., Plotnikov D.E., Sputnikovyi servis monitoringa sostoyaniya rastitel'nosti ("Vega") (Satellite service for vegetation monitoring VEGA), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 190–198.
 7. Palkin N., Oreshnikov V., Biznes-analitika. Ot dannykh k znaniyam (Business-analytics. From data to knowledge), In: *OLAP analiz* (OLAP analysis), St. Petersburg: Piter, 2013, pp. 184–192.
 8. Tolpin V. A., Balashov I.V., Efremov V. Yu., Loupian E.A., Proshin A. A., Uvarov I.A., Flitman E. V., Sozdanie interfeisov dlya raboty s dannymi sovremennykh sistem distantsionnogo monitoringa (sistema GEOSMIS) (The GEOSMIS system: developing interfaces to operate data in modern remote monitoring systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 93–108.
 9. Moss L.T., Atre S., *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*, Addison-Wesley, 2003. 576 p.