

УДК 004.67:[528.88+910.27]

РАЗВИТИЕ СЕТЕВЫХ СЕРВИСОВ ГЕОПОРТАЛА СПУТНИКОВОГО РАДИОТЕПЛОВИДЕНИЯ: ПРОЕКТ ИКАР

Д.М. Ермаков^{1,2}, А.П. Чернушич¹

¹Фрязинский филиал Института радиотехники и электроники им. В.А. Котельникова Российской академии наук (РАН);

²Институт космических исследований РАН

dima@ire.rssi.ru, andrey@fireras.su

Аннотация

Обсужден прогресс в развитии сетевых сервисов и технологий ранее созданного авторами геопортала спутникового радиотепловидения. Исходной задачей геопортала было обеспечить потенциальных пользователей доступом к результатам пространственно-временной интерполяции геофизических атмосферных полей. Практическая реализация потребовала внедрения технологии динамической генерации продуктов обработки. Использование этой технологии расширило круг решаемых задач и привело к созданию специализированного сервиса виртуальной интеграции пространственных геоданных ИКАР (Интерактивного Калькулятора для Атмосферных Расчетов). В работе описаны интерфейсные и технологические принципы, заложенные в основу проекта ИКАР, рассмотрены некоторые детали программно-алгоритмической реализации, приведены примеры практического применения.

Ключевые слова: геопортал, сетевые сервисы, спутниковое радиотепловидение, интерфейсы, программно-алгоритмическая реализация

ВВЕДЕНИЕ

Развитие подхода спутникового радиотепловидения [1, 2] сделало возможным массовую обработку открытой информации спутникового радиотеплового мониторинга для получения глобальных анимированных полей геофизических параметров системы океан–атмосфера на многолетних интервалах наблюдений.

Результаты такой обработки представляют существенный интерес как с точки зрения визуализации и анализа динамики атмосферных процессов, включая экстремальные погодные явления и атмосферные катастрофы [3–5], так и для повышения эффективности дальнейшего совместного анализа с данными других наблюдений за счет более точного взаимного пространственно-временного совмещения.

В 2015 году авторами подхода был запущен геопортал спутникового радиотепловидения [6], одной из главных целей которого является обеспечение свободного доступа потенциальных пользователей к результатам интерполяционной обработки спутниковых данных. В настоящий момент основные продукты обработки включают поля интегрального влагосодержания атмосферы (TPW), полного водозапаса облаков (CLW) и приводного ветра (WND), генерируемые на глобальной сетке с шагом 0,25°. Опорная коллекция охватывает интервал наблюдений с 2004 по 2015 гг. и будет расширяться по мере накопления спутниковой информации. Пользователь имеет возможность просмотра всех данных в режиме анимации, а также заказа полей, интерполированных к любому заданному моменту времени в интервале наблюдений (<http://fire.fryazino.net/tpw/Fields.aspx>).

Принципиальной особенностью современной архитектуры геопортала является динамическая генерация продуктов по запросу пользователя [6–8]. Для геопортала спутникового радиотепловидения эта особенность исходно обусловлена физическими ограничениями на объем хранимой информации при одновременной необходимости обслуживания пользовательских заказов с произвольно задаваемыми параметрами привязки по времени. Дальнейший анализ реализованной схемы работы показал возможность ее эффективного обобщения, при котором заказ данных можно рассматривать как частный случай операции над ними. Выработанное в результате интерфейсное и программно-алгоритмическое решение привело к созданию и развитию проекта ИКАР, предлагающему новые подходы к организации распределенной обработки виртуально интегрированных геоданных из различных открытых источников (специализированных сетевых архивов и баз спутниковых данных, геопорталов и т. д.).

Настоящая работа описывает интерфейсные и технологические принципы, заложенные в основу проекта ИКАР, с углубленным рассмотрением отдельных

деталей программно-алгоритмической реализации, а также примерами практического использования. Существенной проблемой традиционных интерфейсных и программных решений, применяемых при организации геопорталов, является обеспечение функциональной масштабируемости [9, 10]. Наибольшее распространение получил подход, ограничивающий возможности обработки данных некоторым фиксированным набором инструментов. Благодаря этому ограничению удается, с одной стороны, гарантировать корректную работу с исходным и расширяемым в процессе конвейерной обработки множеством данных, а с другой – обеспечить каждый из инструментов удобным графическим пользовательским интерфейсом. Однако такое ограничение представляется излишне жестким. Оно препятствует не только введению в анализ собственных процедур обработки, но и, в ряде случаев, адекватной настройке имеющихся.

Особенно остро эта проблема проявляется при наличии сложных узкоспециализированных процедур, представляющих наибольший интерес при глубоком тематическом анализе данных. Как правило, декомпозиция таких процедур сложна и/или неэффективна, так как результаты промежуточной обработки не имеют смысла вне контекста конкретной задачи. С другой стороны, по мере усложнения алгоритма обработки число его потенциально востребованных модификаций возрастает почти неограниченно, что резко снижает эффективность применения стандартных интерфейсов. Возможный подход к преодолению очерченной проблемы предложен в проекте ИКАР, ориентированном на комплексную тематическую обработку распределенных геоданных.

ИНТЕРФЕЙСНОЕ РЕШЕНИЕ

Проект ИКАР нацелен на организацию распределенных вычислений виртуально интегрированных данных дистанционных наблюдений Земли в интересах решения задач тематической обработки. С этой точки зрения одним из приоритетов проекта была разработка достаточно гибкого интерфейса, позволяющего описывать и модифицировать не только стандартные инструменты анализа больших массивов данных (срезы, подмножества, гистограммы распределений, статистики и т. п.), но и сложные связи между различными типами данных, представляемые в аналитическом виде.

Распространенные графические пользовательские интерфейсы, как указано

выше, не обеспечивают необходимой степени гибкости. Универсальным представлением алгоритма из аналитически описанных шагов расчета является программный код на некотором языке программирования. Класс решений, рассматривающий в качестве объекта сетевого обмена не только данные, но и процедуры их разработки, соответствует наиболее современным тенденциям развития технологий распределенной обработки информации, см., например, [8, 11] и цитируемую в них литературу. Однако при разработке проекта ИКАР авторы стремились к максимально простому в применении решению, не требующему от пользователя навыков программирования. В итоге была предложена и реализована идея калькулятора над таблично описанными данными. В основу было положено предположение о том, что все пространственные данные, потенциально используемые в проекте ИКАР, имеют привязку по времени и могут быть без существенной потери точности и информативности интерполированы на общую расчетную сетку. Сетка имеет 1440 узлов по горизонтали и 720 узлов по вертикали и дает полное покрытие Земли с постоянным шагом между узлами $0,25^\circ$. Вычисления, задаваемые вводимой в текстовом виде формулой, производятся независимо в каждом узле сетки. В результате формируется массив значений, определенный на той же расчетной сетке. Для удобства визуализации итоговые значения с помощью линейного преобразования приводятся к диапазону от 0 до 250 и округляются до целых, после чего отображаются на экране в виде растрового изображения с заданной цветовой палитрой и могут быть сохранены на компьютере пользователя в файл bmp стандартного формата. Коэффициенты калибровки сохраняются в том же файле, как подробно описано в документации на странице проекта ИКАР [12], для возможности обратного пересчета к калиброванным значениям.

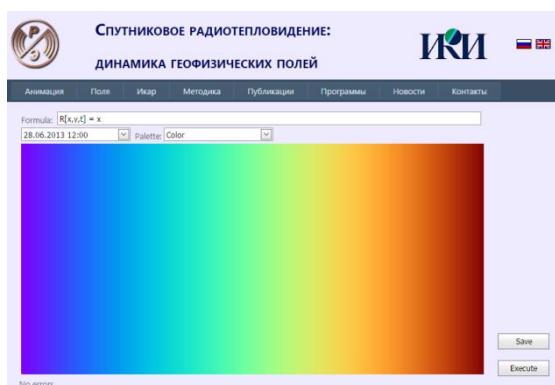
Таким образом, как и в стандартном калькуляторе, реализована возможность организации вычислений путем их аналитического описания в виде текстовой строки (набора строк). Первым принципиальным расширением базового функционала является возможность введения в выражение наряду с константами переменных двух типов – простых скалярных переменных (например, содержащих номера строк, метки времени и т. п.) и трехмерных массивов (например, по-

лей геофизических параметров на заданный момент времени). Проиллюстрируем эту возможность на двух элементарных примерах.

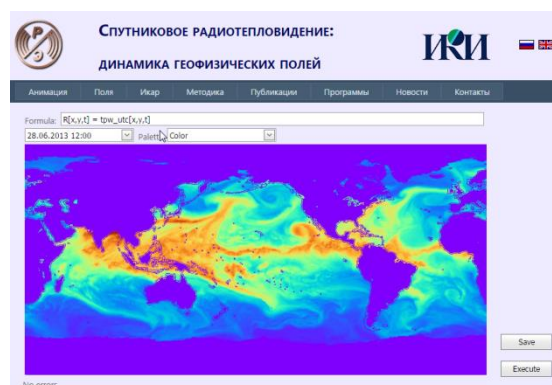
В качестве первого примера рассмотрим формулу с простой скалярной переменной. Ее запись в ИКАР выглядит следующим образом:

$$R[x, y, t] = x, \quad (1)$$

где R – имя результата (некоторая разрешенная комбинация букв и цифр); квадратные скобки означают, что предшествующая переменная описывает массив (результат определен во всех узлах сетки для расчетного момента времени); x – номер столбца (горизонтальная координата узлов сетки), y – номер строки (вертикальная координата узлов сетки), t – расчетный момент времени (задается пользователем с помощью интерфейсного элемента типа стандартного календаря и исчисляется в сутках и долях суток от полуночи 1.01.1990). Формула (1) приравнивает результат в каждом узле сетки горизонтальной координате соответствующего узла. После перенормировки значения R монотонно растут по горизонтали от 0 до 250, так что на экране фактически будет визуализирована используемая цветовая шкала (рис. 1а).



а)



б)

Рис. 1. Визуализация расчета а) по формуле (1); б) по формуле (2)

В качестве второго примера рассмотрим формулу с трехмерным массивом:

$$R[x, y, t] = tpw_utc[x, y, t]. \quad (2)$$

В данном случае значение R в каждом узле совпадает со значением интегрального влагосодержания атмосферы, вычисленного в том же узле и для заданного пользователем момента времени (рис. 1б). Имя массива (в данном случае `trw_utc`) уникально для каждого типа продукта, который, в свою очередь, определяется не только типом геофизического параметра, но и происхождением (источником данных), способом пространственно-временной привязки и т. д. Подробнее этот вопрос рассмотрен в разделе «Продукты и источники данных».

Другими важными расширениями синтаксиса выражений в ИКАР являются условный оператор (оператор ветвления) и возможность определения пользователем новых функций. Оператор ветвления соответствует широко применяемой записи с фигурной скобкой, при которой процесс вычисления конечного результата управляется одним или несколькими условиями. Так, например, приведенная ниже формула в синтаксисе ИКАР

$$R[x, y, t] = (x > y) ? x : y \quad (3)$$

означает

$$R[x, y, t] = \begin{cases} x, & x > y \\ y, & x \leq y \end{cases} \quad (4)$$

Возможность введения определяемых пользователем функций является стандартным способом расширения синтаксиса выражений, направленным на упрощение записи сложных выражений, в том числе, неоднократно использующих однотипную последовательность вычислений. В качестве примера рассмотрим следующую формулу:

$$R[x, y, t] = (d(x, y, t) > 0.2) ? d(x, y, t) : 0; \quad d(x, y, t) = \text{rss_sst}[x, y, t] - \text{rss_sst}[x, y, t - 100] \quad (5)$$

Результат вычислений равен разности между температурами поверхности океана (`rss_sst`) в заданную дату и за 100 суток до нее в той же точке, если эта разность превосходит $0,2^\circ\text{C}$, а в противном случае приравнивается 0. Для компактности введена функция разности температур $d(x, y, t)$, которая сначала используется в выражении и лишь затем определяется в явном виде (после точки с запятой). Рис. 2а иллюстрирует результат вычислений по формуле (5) при заданной

текущей дате 28.08.2013. Пользователь имеет возможность изменить дату с помощью стандартного интерфейсного элемента типа календаря, чем обеспечивается возможность применения одной и той же расчетной формулы для произвольного момента времени (при доступности соответствующей спутниковой или иной входной информации). Рис. 2б иллюстрирует новый результат вычислений по (5) при смене даты на 28.02.2013.

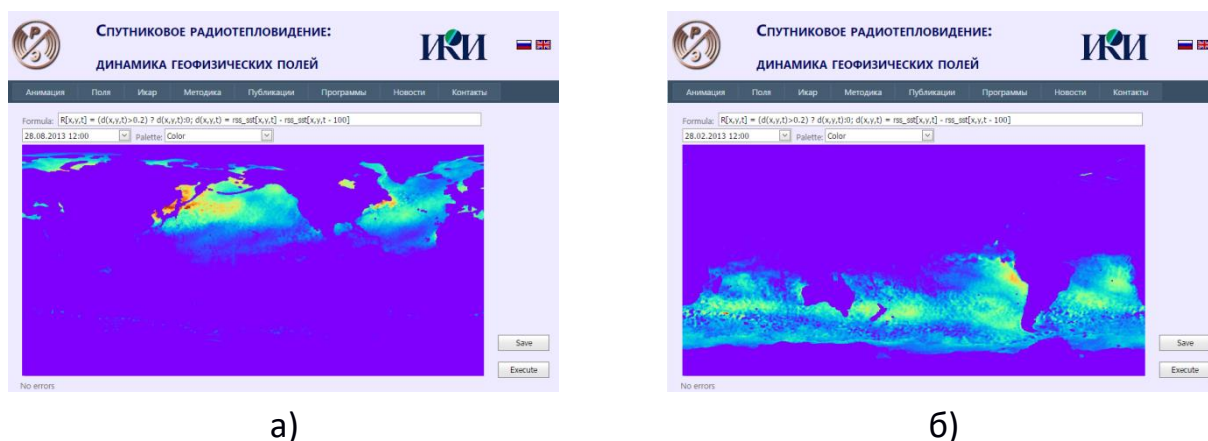


Рис. 2. Визуализация расчета (5) для дат: а) 28.08.2013; б) 28.02.2013

В следующем разделе приведено формальное описание класса текстовых строк, интерпретируемых ИКАР.

СИНТАКСИС И СЕМАНТИКА ИКАР

Синтаксис и семантику интерпретируемых в ИКАР текстовых строк удобно объединить на основе синтаксических графов (диаграмм Вирта) [13]. При этом для компактности изложения предварительно оговорим некоторые основные особенности.

Интерпретируемый ИКАР текст (последовательность символов) состоит из основной *формулы*, связывающей результат с *выражением* и, возможно, одним или несколькими определениями *функций*, каждое из которых аналогично основной формуле, но вместо символического обозначения результата начинается символическим обозначением ранее использованной в выражениях функции. Каждое новое определение функции отделено от предшествующего текста точкой с запятой. Таким образом, весь текст состоит из набора строк, отделенных друг от друга точкой с запятой, причем первая строка является *формулой*, а

остальные (при наличии) являются декларациями *функций*. Поэтому достаточно привести полные описания синтаксиса и семантики *формулы* и *функции* по отдельности.

Далее будем считать заранее определенными понятие *имени* как последовательности из латинских букв и цифр от 0 до 9, начинающейся с буквы, а также стандартного десятичного представления числовых констант с плавающей точкой. Будем также считать все пробельные символы необязательными разделителями лексем, допустимыми в тексте и игнорируемыми при синтаксическом разборе. Запрещен разрыв пробельными символами лексемы (имени, числа или операции, представленных несколькими символами), а также разрыв между именем функции (массива) и последующей открывающей круглой (квадратной) скобкой. Последнее требование связано с тем, что неразрывное сочетание имени и скобки принято в качестве синтаксического признака функции и массива в отсутствие механизма их предварительной декларации.

Диаграмма на рис. 3 дает полное описание строки, соответствующее понятию *формулы* в ИКАР. По способу, предложенному Виртом, уточнение идет в направлении от общего к частному, при этом элементарные термины (лексемы) представлены в овальных рамках, а составные – в прямоугольных.

Как показано на рис. 3, в формуле последовательно перечисляются: символическое имя результата, его *параметры* в квадратных скобках и собственно вычисляемое *выражение* после знака «равно». *Параметры* содержат имена скалярных переменных, соответствующих номеру столбца (номеру узла в строке), номеру строки и моменту времени. Для первых двух переменных допускается также задание диапазона изменений (минимального и максимального значений). По умолчанию диапазон изменений охватывает всю расчетную сетку. Имена переменных назначаются пользователем произвольно и могут быть в дальнейшем использованы в вычисляемых выражениях.

В свою очередь, *выражение* состоит из одной *части* либо из двух *частей*, объединенных логическим *условием*: «равно», «не равно», «меньше», «больше», «меньше или равно», «больше или равно». В последнем случае выражение считается равным единице, если записанное в нем условие выполняется, и нулем, если условие не выполняется.

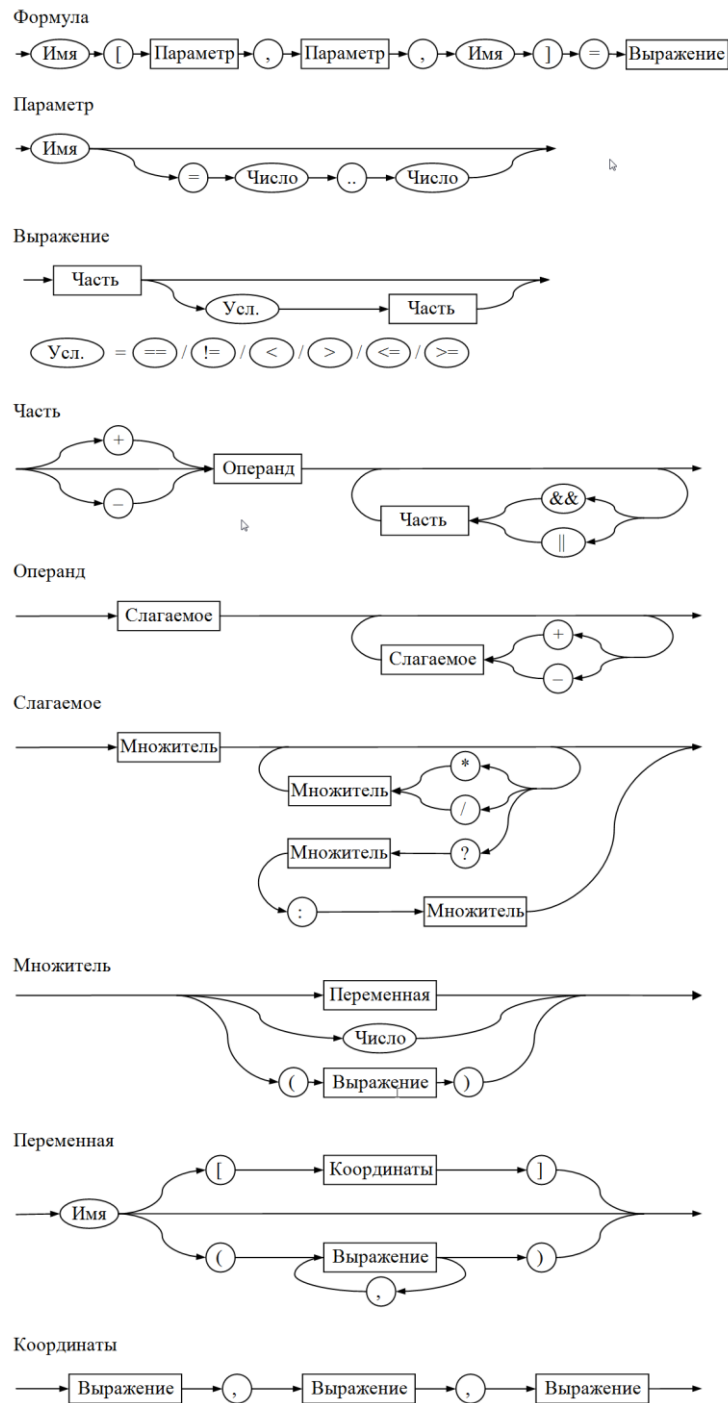


Рис. 3. Синтаксический граф для понятия «формула» в ИКАР

Часть выражения определяется рекурсивно. Она может состоять из одного операнда (возможно, с предшествующим ему знаком «плюс» или «минус») или объединяться с другими частями с помощью логических операторов конъюнк-

ции («И», «&&») и/или дизъюнкции («ИЛИ», «| |»). Синтаксический разбор составных частей осуществляется до их полной декомпозиции на операнды, объединенные логическими операциями.

Операнд представляет собой *слагаемое* или несколько *слагаемых*, связанных операторами сложения и/или вычитания.

Слагаемое представляет собой *множитель* или несколько множителей, объединенных операторами умножения и/или деления, и может быть дополнительно завершено операцией ветвления, состоящей из символа «?», *множителя*, символа «:» и *множителя* альтернативы. Правило вычисления результата в операции ветвления таково: если значение множителя, предшествующего знаку «?», отлично от нуля, то результатом является значение *множителя*, непосредственно следующего за знаком «?»; в противном случае результатом является значение *множителя*, следующего за знаком «:» (альтернативы), см. также описание к формулам (3) и (5).

Множитель может представлять собой *переменную*, число или новое *выражение* в круглых скобках.

Переменная может быть представлена одним *именем* (соответствует простой скалярной величине), либо *именем с координатами* в квадратных скобках (соответствует трехмерному массиву), либо *именем с выражением* или несколькими перечисленными через запятую *выражениями* в круглых скобках (соответствует функции одной или нескольких переменных).

Координаты представляют собой последовательность из трех *выражений*, разделенных запятыми.

Таким образом, синтаксический разбор формулы по правилам, изображенным на схеме рис. 3, позволяет либо представить конечный результат в виде композиции логических и арифметических операций над значениями переменных, массивов и функций, либо выявить в формуле синтаксическую ошибку. Приоритет операций, в основном, также определен указанными синтаксическими правилами: наибольший – у операций деления и умножения, ниже – у операций сложения и вычитания, еще ниже – у логических операций конъюнкции и дизъюнкции (дополнительно принято, что приоритет первой из них выше), самый низкий

– у операций сравнения. Операции с равным приоритетом выполняются в порядке перечисления в выражении (слева направо).

Для замыкания вычислительной схемы необходимо дополнительно определить способ получения значений входных данных – простых переменных, массивов и функций.

Принято, что в *выражениях* ИКАР могут фигурировать только простые переменные, соответствующие параметрам результата (номеру узла, номеру строки и момента времени) или формальным параметрам функции (внутри ее определения). При этом в ходе вычислений значения номера узла и номера строки последовательно пробегают по диапазонам разрешенных значений, и для каждого сочетания номеров узла и строки производится расчет по формуле. Момент времени фиксируется до начала вычислений и может быть установлен пользователем с помощью графического интерфейса ИКАР, как описано выше. Инициализация параметров функции происходит перед ее вызовом из вычисляемого выражения. Таким образом, в любой момент вычислений значения всех простых переменных однозначно определены.

Массивы в ИКАР представляют собой семейства матриц, параметризованных моментом времени. Для установленного пользователем момента времени каждый массив является набором значений, определенных во всех узлах расчетной сетки 1440 (столбцов) x 720 (строк). При установке нового момента времени значения массива, вообще говоря, требуют обновления. Массив представляет собой удобный способ описания некоторого двумерного поля (поля геофизического параметра), меняющегося во времени. Для указания на конкретный геофизический параметр используется уникальное имя массива. Таким образом, в формулах ИКАР разрешено использование массивов с именами из ограниченного, заранее определенного набора, каждое из которых соответствует определенному продукту. Продукт может либо храниться на сервере геопортала спутникового радиотепловидения, либо быть результатом динамической генерации и/или виртуальной интеграции данных других архивов, баз спутниковых данных или геопорталов. Для обеспечения единообразного включения в формулы любых типов продуктов реализован унифицированный инструмент загрузки продуктов в процессе их использования в расчетах. Список продуктов на данный момент содержит 7

наименований (см. ниже) и будет расширяться по мере развития геопортала.

Проект ИКАР допускает использование двух типов функций: библиотечных и пользовательских. Библиотечные функции имеют заранее определенные имена и фиксированное (вообще говоря, различное для разных функций) число параметров. В настоящий момент определен набор из 17 функций, содержащий, в том числе, наиболее употребительные тригонометрические и показательные функции, а также ряд других. Полный набор описан на сайте проекта по адресу <http://fire.fryazino.net/tpw/AboutIcar.aspx>.

Пользовательские функции являются одним из традиционных способов расширения синтаксиса для более компактного и удобного описания процедуры расчета, а также для организации нетривиальных способов вычислений, например, рекурсии. Особенностью синтаксиса ИКАР является то, что допускается употребление функций в вычисляемых выражениях до ее формального определения. Безусловным требованием является строгое соответствие вызова функции ее определению (т. е. точное совпадение имени функции и числа параметров). В качестве примера см. описание формулы (5).

Синтаксис определения функции иллюстрирует рис. 4.

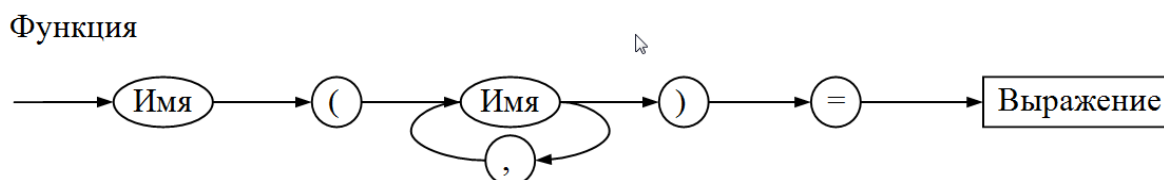


Рис. 4. Синтаксический граф определения функции в ИКАР

Согласно рис. 4 в описании *функции* последовательно перечисляются: (уникальное) *имя* функции, набор имен формальных параметров, разделенных запятыми, внутри круглых скобок и вычисляемое *выражение* после знака «равно». Определение *выражения* дано выше при описании синтаксиса *формулы* (рис. 3).

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Программная реализация ИКАР построена на принципе частичной компиляции. После ввода пользователем текста основной формулы и определения

функций и запуска вычислений сначала производится однократный синтаксический разбор текста. Разбор включает проверку на формальное соответствие правилам синтаксиса ИКАР, построение таблиц имен с определением типа каждой новой переменной (имя результата, имя переменной столбца, строки и момента времени, имя массива, имя функции, имя формального параметра функции) и присвоением каждой из них уникального целочисленного идентификатора. Для всех обнаруженных массивов осуществляется проверка принадлежности их имен заранее определенному набору (каждому уникальному имени соответствует свой загрузчик данных). Для имени каждой обнаруженной функции осуществляется проверка наличия соответствующего пользовательского определения или принадлежности к набору библиотечных функций. Для пользовательских функций строятся *сигнатуры*, состоящие из уникального имени функции, знака «@» и числа, равного числу параметров функции. Кроме того, автоматически производится проверка непротиворечивости всех деклараций (отсутствие совпадений имен, относящихся к переменным разного типа, или к нескольким разным функциям). Далее для основной формулы и каждого определения функции формируется по паре стеков, состоящей из стека операций и стека данных.

Алгоритм формирования стеков вполне аналогичен алгоритму перевода обычного скобочного выражения (в инфиксной записи) в бесскобочную форму (префиксную или постфиксную) [14]. В данном случае операторы и операнды разносятся в разные стеки (стек операций и стек данных). Благодаря введению двух стеков можно считать все данные числами с плавающей запятой, представленными восьмью байтами, а все операции кодировать четырехбайтовыми целочисленными идентификаторами. При этом к списку операторов, определенных синтаксисом ИКАР, необходимо добавить снятие числа со стека данных («&») и обращение к таблице имен («\$»).

В качестве примера рассмотрим инфиксное, постфиксное и стековое представления простого арифметического выражения, рис. 5.

$(2+3) \cdot 5$	$2\ 3\ +\ 5\ *$	стек данных: 2 3 5;	стек операций:
			& & + & *
а)	б)	в)	

Рис. 5. Выражение в инфиксном (а), постфиксном (б) и стековом (в) представлениях; вершины стеков – справа

На этом примере разберем алгоритм вычисления выражений. Алгоритм устроен рекурсивно. На первом шаге снимается код операции с текущей вершины стека операций. В примере, приведенном на рис. 5, это умножение, «*». Операция умножения требует двух операндов (сомножителей), каждый из которых является результатом исполнения последующих операций. Поэтому процедура вычислений применяется рекурсивно к следующей операции на вершине стека операций, в данном случае, к операции снятия числа со стека данных, «&». Операция «&» снимает число 5 с вершины стека данных и возвращает его в качестве своего результата – первого множителя. Второй множитель является результатом выполнения операции сложения, «+». Эта операция также требует двух операндов – слагаемых. Каждое из слагаемых, по аналогии с вышеприведенным описанием, является результатом операции снятия данных со стека, которые последовательно возвращают числа 3 и 2 (при исполнении операций, в которых порядок операндов является существенным, следует учесть, что со стека данных они снимаются в порядке, противоположном их появлению в исходном выражении). После этого для операции «+» оказываются вычисленными оба слагаемых, и она возвращает результат (их сумму) в качестве второго множителя для операции «*». В свою очередь, операция «*» получает оба операнда и возвращает результат их перемножения. Поскольку рекурсивное вычисление окончено и стек операций исчерпан (необходимое условие синтаксической корректности), последний результат является значением всего выражения.

При наличии в выражении переменных в стеке операций используется также операция обращения к таблице имен, а в стеке данных – идентификаторы соответствующих переменных. Разберем это на примере приведенной выше формулы (2). Сформированная по ней таблица имен схематично представлена в Таблице 1.

Как указано выше, таблица имен формируется в процессе синтаксического разбора. По соглашению, имени результата присваивается идентификатор 0, именам параметров (моменту времени и номерам столбца и строки) – отрицательные

идентификаторы. Остальным именам присваиваются положительные идентификаторы, возрастающие в порядке появления имени в выражении. При повторном возникновении некоторого имени в выражении осуществляется проверка того, что синтаксически имя описывает тот же тип, что и раньше, а в случае функций – что функция имеет одну и ту же сигнатуру (то же число параметров).

Таблица 1. Схема таблицы имен для расчетной формулы (2)

Имя	Идентификатор	Тип
y	-3	Простая переменная (номер строки)
x	-2	Простая переменная (номер столбца)
t	-1	Простая переменная (момент времени)
R	0	Имя результата
tpw_utc	1	Трехмерный массив (поле геопараметра)

По выражению формулы (2) с учетом значений идентификаторов из Таблицы 1 будут сформированы стеки операций и данных, представленные в Таблице 2.

Таблица 2. Схема стеков для расчетной формулы (2)

Стек	Содержимое
данных:	$-2 -3 -1 1$
операций:	$\$ \$ \$ \$$

Как и в других случаях, исполнение начинается снятием идентификатора операции с вершины стека операций, в данном случае, операции обращения к таблице имен, «\$». Операция снимает число с вершины стека данных, 1, и интерпретирует его как идентификатор в таблице имен. Согласно Таблице 1 идентификатор 1 относится к массиву, поэтому для вычисления текущего значения необходимо получить три индекса (по числу измерений массива). Первый индекс равен результату следующей операции «\$», которая снимает число -1 со стека данных и интерпретирует его как идентификатор в таблице имен. Этот идентификатор относится к простой переменной (момент времени), и для получения ее значения операция конвертирует идентификатор в адрес в области временной памяти, где хранятся текущие значения всех простых переменных из таблицы имен. Процесс

повторяется со вторым и третьим индексом, после чего начальная операция получает значения всех трех индексов массива и вызывает загрузчик данных для получения значения элемента массива по вычисленным индексам. Работа загрузчика данных описана отдельно.

Вычисление значения функции осуществляется аналогично. Однако разные функции могут иметь разное число аргументов. Поэтому операция вычисления функции дополнительно получает по имени функции ее сигнатуру, определяет число аргументов, которые должны быть вычислены перед вызовом функции.

Итак, частичная компиляция состоит в переводе главной формулы и определений функций в стековое представление и формировании таблиц имен. Исполнение осуществляется с помощью программных интерпретаторов, работающих по принципу стековых машин. В начале исполнения организуется два вложенных цикла. Внешний цикл меняет текущее значение номера строки (второго аргумента результата) от минимального значения до максимального. Внутренний цикл меняет значения номера столбца. Для каждого сочетания номеров столбца и строки осуществляется вызов программного интерпретатора, относящегося к главной формуле. Во временную память интерпретатора, предназначенную для хранения параметров (номеров строки и столбца и момента времени), заносятся их текущие значения. Программный интерпретатор делает копии своих стеков операций и данных (т. к. они будут меняться в ходе исполнения и должны быть восстановлены при вызове со следующим набором параметров). Затем программный интерпретатор снимает первую операцию с вершины стека операций и начинает рекурсивное вычисление результата, как описано выше. Вычисленное значение заносится в элемент массива во временной памяти, отведенный для хранения результата. Индекс элемента определяется текущими номерами столбца и строки.

При достижении интерпретатором операции вызова пользовательской функции исполнение продолжается рекурсивно. Сначала определяется число аргументов и осуществляется вычисление их значений, как описано выше. Далее управление передается новому программному интерпретатору, работающему с копиями стеков операций и данных (и таблицей имен), построенных для соответ-

ствующей функции. При этом временная память интерпретатора инициализируется вычисленными значениями параметров функции. При получении окончательного результата управление возвращается программному интерпретатору, инициировавшему вызов функции. Программный интерпретатор библиотечной функции не работает со стеками. Процедура вычисления результата в этом случае жестко зафиксирована в программном коде, и интерпретатор вызывает ее с переданными ему значениями параметров.

При достижении интерпретатором операции получения значения массива происходит вызов соответствующего *загрузчика* данных. В начале работы загрузчик проверяет, не были ли запрашиваемые данные уже загружены ранее. При первом запросе загрузчик вызывает процедуру загрузки, специфическую для каждого типа данных. В общем случае, загрузка может состоять в установлении связи с удаленным сервером по одному из стандартных протоколов сетевого обмена (ftp, http), получении требуемых файлов данных, их распаковке, декодировании и пространственно-временной привязке с интерполяцией на общую расчетную сетку. Таким образом, процедура запроса данных может составлять существенную часть времени исполнения, при этом она выполняется многократно (как минимум, по одному разу для каждого возможного сочетания индексов узлов сетки). Поэтому все загружаемые в рамках одного сеанса работы данные временно сохраняются на сервере геопортала и сразу готовы к работе при повторном запросе. Проверка готовности запрашиваемых данных состоит в построении уникального идентификатора данных по имени массива (типу данных) и текущему моменту времени и в поиске адреса во временной памяти, ассоциированного с этим идентификатором. Если такой адрес обнаружен, загрузчик сразу обращается к данным по этому адресу. В противном случае загрузчик выполняет полную процедуру загрузки, строит новый идентификатор данных и связывает его с областью памяти, отведенной под вновь загруженные данные.

ОБЕСПЕЧЕНИЕ ДАННЫМИ И ФУНКЦИОНАЛЬНОЕ НАПОЛНЕНИЕ

Как указано выше, в настоящее время в выражениях ИКАР могут быть использованы массивы данных семи различных типов. Они представлены в Таблице 3.

Таблица 3. Типы и источники используемых данных

Имя	Описание	Источник
tpw_ltw	Интегральное влагосодержание атмосферы при заданном местном времени	ГСР ¹
tpw_utc	Интегральное влагосодержание атмосферы при заданном всемирном времени	ГСР
clw_ltw	Полный водозапас облаков при заданном местном времени	ГСР
clw_utc	Полный водозапас облаков при заданном всемирном времени	ГСР
wnd_ltw	Скорость приводного ветра при заданном местном времени	ГСР
wnd_utc	Скорость приводного ветра при заданном всемирном времени	ГСР
rss_sst	Температура поверхности океана в местный полдень	RSS ²

¹ Геопортал спутникового радиотепловидения (<http://fire.fryazino.net/tpw/>)

² Архив Remote Sensing Systems (<http://www.remss.com/>)

Основным источников данных для расчетов в ИКАР в настоящее время является геопортал спутникового радиотепловидения. Однако, благодаря унифицированной архитектуре загрузчиков, обеспечена возможность использования данных из других архивов, баз данных и геопорталов по принципу виртуальной интеграции. Эта возможность проиллюстрирована на примере полей температуры поверхности океана, получаемых непосредственно из архива компании Remote Sensing Systems (США). Авторы рассматривают различные варианты расширения номенклатуры виртуально интегрируемых продуктов, производимых отечественными и зарубежными специалистами.

Безусловно, важным аспектом проекта является обеспечение широкого набора библиотечных аналитических функций. В настоящее время, как указано выше, реализовано 17 библиотечных функций. Они перечислены в Таблице 4.

Таблица 4. Библиотечные функции ИКАР

Функция	Аргументы	Значение
$\sin(a)$	Угол в радианах	Синус аргумента
$\cos(a)$	Угол в радианах	Косинус аргумента
$\tan(a)$	Угол в радианах	Тангенс аргумента
$\text{asin}(a)$	Безразмерный аргумент	Арксинус в радианах
$\text{acos}(a)$	Безразмерный аргумент	Арккосинус в радианах
$\text{atan2}(y,x)$	y – противолежащий катет, x – прилежащий катет	Арктангенс в радианах
$\exp(a)$	Безразмерный аргумент	Экспонента аргумента
$\log(a)$	Безразмерный аргумент	Натуральный логарифм
$\log_{10}(a)$	Безразмерный аргумент	Десятичный логарифм
$\text{pow}(a,b)$	Безразмерные аргументы	a в степени b
$\text{lat}(y)$	Безразмерный аргумент	Широта строки y
$\text{lon}(x)$	Безразмерный аргумент	Долгота столбца x
$\text{time}(Y,M,D,h,m,s)$	Год, месяц, число, час, минута, секунда	Промежуток времени в сутках и долях суток от полуночи 01.01.1990 до указанного момента
$\text{rad}(a)$	Угол в градусах	Радианная мера угла
$\text{min}(a,b)$	Безразмерные аргументы	Наименьшее из a и b
$\text{max}(a,b)$	Безразмерные аргументы	Наибольшее из a и b
$\text{abs}(a)$	Безразмерный аргумент	Модуль аргумента

Возможность внедрения новых библиотечных функций обеспечена архитектурой программного интерпретатора. Добавление функции осуществляется путем интеграции в систему нового специфичного программного интерпретатора. Авторы планируют дальнейшее расширение набора библиотечных функций.

ПРАКТИЧЕСКОЕ ИСПОЛЬЗОВАНИЕ

Универсальность подхода, заложенного в основу проекта ИКАР, проявляется в возможности организации вычислений как характерных для большинства

геопорталов (различные статистические характеристики, профили, срезы данных, временные ряды и т. п.), так и специфических, направленных на решение конкретных практических задач. Важно, что в последнем случае сохраняется необходимая гибкость, позволяющая легко вносить произвольные модификации в расчетные выражения в интерактивном режиме, практически сразу отслеживая и сопоставляя результаты внесенных изменений. Таким образом, проект ИКАР может быть использован как для первичного анализа данных, так и для разработки, модификации и отладки алгоритма их обработки, а также для организации потоковой обработки больших массивов данных с использованием развитого алгоритма.

Для эффективной потоковой обработки на геопортале спутникового радиотепловидения реализована сетевая служба (веб-сервис), открывающая функциональный интерфейс к ИКАР. Функция передает на сервер расчетную формулу и расчетное время, а получает результат в виде массива на расчетной сетке или текстовое сообщение об ошибке. Таким образом, обеспечена возможность разработки клиентского приложения, выполняющего потоковую обработку с использованием функционала программного ядра ИКАР.

Важно отметить особенность сетевого обмена данными при реализации вычислений в ИКАР, состоящую в том, что сетевой трафик между компьютером пользователя и сервером геопортала минимизирован. Пользователь передает на сервер только текст основной формулы и определения функций и получает только массив расчетных значений или текстовое сообщение об ошибке. Виртуальная интеграция данных осуществляется путем сетевого обмена между сервером геопортала и распределенными архивами и не влияет на объем трафика, передаваемого пользователю. Также важно, что передаваемый пользователем текст является одновременно и способом документации (описания и сохранения) алгоритма обработки. Этот текст может быть записан в текстовый файл и впоследствии использован повторно как в интерактивном сеансе работы, так и с помощью программы-клиента.

В качестве иллюстрации практического использования ИКАР приведем пример расчета так называемого ТВ-критерия [15]. В синтаксисе ИКАР он может быть записан следующим образом:

$$R[x, y = 240..480, t] = \max(0, 100 \cdot \text{del}(\text{tpw_ltw}[x, y, t]/10, \text{rss_sst}[x, y, t])) \cdot \text{abs}(\sin(\text{rad}(\text{lat}(y)))) ;$$

$$\text{del}(A, B) = (A - 5,07)/0,38 - (B - 27)/4 \quad (6)$$

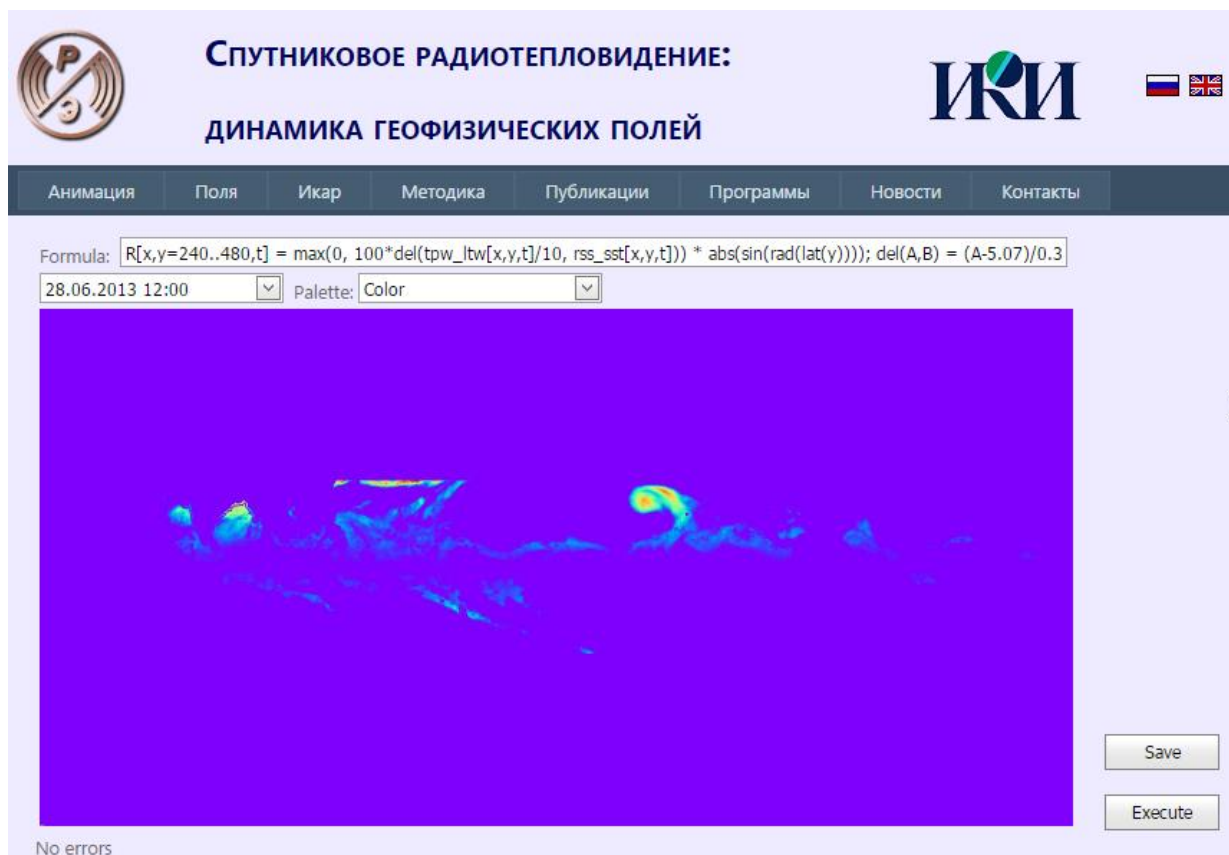


Рис. 6. Расчет по формуле (6) для даты 28.06.2013

На рис. 6 показан результат расчета для 28.06.2013. Условное время – местный полдень (к этому времени привязаны данные температуры поверхности океана). Область расчета ограничена низкими широтами (от 30° с.ш. до 30° ю.ш.), что повышает вычислительную эффективность и предотвращает возникновение дополнительных максимумов ТВ-критерия, связанных с ростом силы Кориолиса в высоких широтах, см. [15]. Реализация потоковой обработки дает возможность широкомасштабного статистического анализа связи приведенного критерия с глобальным тропическим циклогенезом.

ЗАКЛЮЧЕНИЕ

В настоящей работе подробно рассмотрен проект ИКАР, являющийся развитием сетевых сервисов геопортала спутникового радиотепловидения. В проекте реализовано новое программно-интерфейсное решение в организации удаленной обработки распределенных геоданных, основанное на идее интерактивного калькулятора. Показано, что решение обеспечивает необходимую гибкость при проектировании инструментов исследования с привлечением спутниковых данных, включая первичный анализ, разработку, модификацию и отладки алгоритма обработки, а также для организации потоковой обработки больших массивов спутниковой информации. Проанализированы детали технической реализации, обеспечивающие вычислительную эффективность и минимизацию трафика сетевого обмена с пользователем. Обсуждены возможности практического применения ИКАР, проиллюстрированные конкретными примерами обработки с использованием как данных геопортала спутникового радиотепловидения, так и другими, виртуально интегрированными архивами данных. Перспективы развития проекта связаны с обеспечением возможности широкой виртуальной интеграции различных архивов, баз геоданных и геопорталов, с расширением библиотеки встроенных аналитических функций, с совершенствованием пользовательских интерфейсов.

Благодарности

Развитие проекта ИКАР поддержано грантом РФФИ № 15-07-04422-А.

СПИСОК ЛИТЕРАТУРЫ

1. *Ермаков Д.М., Шарков Е.А., Чернушич А.П.* Мультисенсорный алгоритм спутникового радиотепловидения // Исследование Земли из космоса. 2016. № 3. С. 37–46.
2. *Ермаков Д.М., Шарков Е.А., Чернушич А.П.* Спутниковое радиотепловидение на синоптических и климатически значимых масштабах // Исследование Земли из космоса. 2016. № 5. С. 3–9.
3. *Шарков Е.А.* Полярный перенос скрытого тепла мезомасштабными тропосферными системами: анимационный анализ микроволновых спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 170–187.

4. *Ermakov D.M., Sharkov E.A., Chernushich A.P.* Satellite radiothermovision of atmospheric mesoscale processes: case study of tropical cyclones // *ISPRS Archives*. 2015. V. 40. No 7/W3. P. 179–186.

5. *Wimmers A.J., Velden C.S.* Seamless advective blending of total precipitable water retrievals from polar-orbiting satellites // *J. Appl. Meteor. Climatol.* 2011. V. 50. No 5. P. 1024–1036.

6. *Ермаков Д.М., Чернушич А.П., Шарков Е.А.* Геопортал спутникового радиотепловидения: данные, сервисы, перспективы развития // *Современные проблемы дистанционного зондирования Земли из космоса*. 2016. Т. 13. № 3. С. 46–57.

7. *Лупян Е.А., Саворский В.П., Шокин Ю.И., Алексанин А.И., Назиров Р.Р., Недолужко И.В., Панова О.Ю.* Современные подходы и технологии организации работы с данными дистанционного зондирования Земли для решения научных задач // *Современные проблемы дистанционного зондирования Земли из космоса*. 2012. Т. 9. № 5. С. 21–44.

8. *Savorskiy V., Lupyan E., Balashov I., Proshin A., Tolpin V., Ermakov D., Chernushich A., Panova O., Kuznetsov O., Vasilyev V.* Basic technologies of web services framework for research, discovery, and processing the disparate massive Earth observation data from heterogeneous sources // *ISPRS Archives*. 2014. V. 40. No 4. P. 223–228.

9. *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. № 4. С. 93–108.

10. *Resch B., Zimmer B.* User experience design in professional map-based geoportals // *ISPRS Int. J. Geo-Information*. 2013. No 2. P. 1015-1037.

11. *Биктимиров М.Р., Елизаров А.М., Щербаков А.Ю.* Тенденции развития технологий обработки больших данных и инструментария хранения разноформатных данных и аналитики // *Электронные библиотеки*. 2016. Т. 19. № 5. С. 390–407.

12. Геопортал спутникового радиотепловидения. Проект ИКАР: Описание.

<http://fire.fryazino.net/tpw/AboutIcar.aspx>.

13. *Wirth N.* Algorithms + data structures = programs. Englewood Cliffs, NJ, USA: Prentice-Hall, 1976. 400 p.

14. *Pratt T.W.* Programming languages: design and implementation. Englewood Cliffs, NJ, USA: Prentice-Hall, 1975. 530 p.

15. *Ростовцева В.В., Гончаренко И.В.* Оценка влияния тепловой стратификации тропосферы на активность тропического циклогенеза по данным спутниковой микроволновой радиометрии // Исследование Земли из космоса. 2014. № 2. С. 3–17.

DEVELOPMENT OF NETWORK SERVICES OF THE GEOPORTAL OF SATELLITE RADIOTHERMOVISION: ICAR PROJECT

D.M. Ermakov^{1,2}, A.P. Chernushich¹

¹*Kotel'nikov Institute of Radio Engineering and Electronics of Russian Academy of Sciences (RAS), Fryazino Branch;*

²*Space Research Institute of RAS*

dima@ire.rssi.ru, andrey@fireras.su

Abstract

The paper discusses the progress in the development of network services and technologies of the geoportal of satellite radiothermovision created by the authors earlier. The initial task of the geoportal was to provide potential users access to the results of the spatiotemporal interpolation of geophysical atmospheric fields. Practical implementation required the implementation of a technology of dynamic generation of processing products. The use of this technology had expanded the range of tasks, and led to the creation of a specialized service of the virtual integration of spatial geo-data ICAR (Interactive Calculator for Atmospheric Research). The paper describes the interface and technological principles underlying the ICAR project, discusses some details of software and algorithmic implementation, provides examples of practical applications.

Keywords: *geoportal, network services, satellite radiothermovision, interfaces,*

software and algorithmic implementation

REFERENCES

1. *Ermakov D.M., Sharkov E.A., Chernushich A.P.* Mul'tisensornyy algoritm sputnikovogo radioteplovideniya // *Issledovaniye Zemli iz kosmosa*. 2016. No 3. P. 37–46.
2. *Ermakov D.M., Sharkov E.A., Chernushich A.P.* Sputnikovoye radioteplovideniye na sinopticheskikh i klimaticheskikh znachimykh masshtabakh // *Issledovaniye Zemli iz kosmosa*. 2016. No 5. P. 3–9.
3. *Sharkov E.A.* Polyarnyy perenos skrytogo tepla mezomasshtabnymi troposfernymi sistemami: animatsionnyy analiz mikrovolnovykh sputnikovyykh dannykh // *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*. 2015. V. 12. No 5. P. 170–187.
4. *Ermakov D.M., Sharkov E.A., Chernushich A.P.* Satellite radiothermography of atmospheric mesoscale processes: case study of tropical cyclones // *ISPRS Archives*. 2015. V. 40. No 7/W3. P. 179–186.
5. *Wimmers A.J., Velden C.S.* Seamless advective blending of total precipitable water retrievals from polar-orbiting satellites // *J. Appl. Meteor. Climatol.* 2011. V. 50. No 5. P. 1024–1036.
6. *Ermakov D.M., Chernushich A.P., Sharkov E.A.* Geoportal sputnikovogo radioteplovideniya: dannye, servisy, perspektivy razvitiya // *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*. 2016. V. 13. No 3. P. 46–57.
7. *Loupian E.A., Savorskiy V.P., Shokin Yu.I., Aleksanin A.I., Nazirov R.R., Nedolugko I.V., Panova O.Yu.* Sovremennyye podhody i tehnologii organizatsii raboty s dannymi distantsionnogo zondirovaniya Zemli dlya resheniya nauchnykh zadach // *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*. 2012. V. 9. No 5. P. 21–44.
8. *Savorskiy V., Lupyan E., Balashov I., Proshin A., Tolpin V., Ermakov D., Chernushich A., Panova O., Kuznetsov O., Vasilyev V.* Basic technologies of web services framework for research, discovery, and processing the disparate massive Earth observation data from heterogeneous sources // *ISPRS Archives*. 2014. V. 40. No 4. P. 223–228.
9. *Tolpin V.A., Balashov I.V., Efremov V.Yu., Loupian E.A., Proshin A.A., Uvarov I.A., Flitman E.V.* Sozdanie interfeysov dlya raboty s dannymi sovremennykh

system distantsionnogo monitoringa (sistema GEOSMIS) // *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*. 2011. V. 8. No 4. P. 93–108.

10. *Resch B., Zimmer B.* User experience design in professional map-based geo-portals // *ISPRS Int. J. Geo-Information*. 2013. No 2. P. 1015–1037.

11. *Biktimirov M.R., Elizarov A.M., Scherbakov A.Yu.* Tendentsii razvitiya tehnologii obrabotki bol'shih dannyh i instrumentariya hraneniya raznoformatnyh dannyh i analitiki // *Elektronnyye biblioteki*. 2016. V. 19. No 5. P. 390–407.

12. Geoportal of satellite radiothermvision. ICAR project: Help. <http://fire.fryazino.net/tpw/AboutIcar.aspx>.

13. *Wirth N.* Algorithms + data structures = programs. Englewood Cliffs, NJ, USA: Prentice-Hall, 1976. 400 p.

14. *Pratt T.W.* Programming languages: design and implementation. Englewood Cliffs, NJ, USA: Prentice-Hall, 1975. 530 p.

15. *Rostovtseva V.V., Goncharenko I.V.* Otsenka vliyaniya teplovoy stratifikatsii atmosfery na aktivnost' tropicheskogo tsiklogeneza po dannym sputnikovoy mikrovolnovoy radiometrii // *Issledovaniye Zemli iz kosmosa*. 2014. No 2. P. 3–17.

СВЕДЕНИЯ ОБ АВТОРАХ

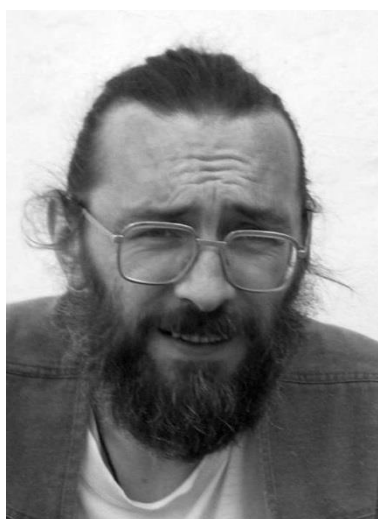


ЕРМАКОВ Дмитрий Михайлович – кандидат физико-математических наук, ведущий научный сотрудник Фрязинского филиала Института радиотехники и электроники им. В.А. Котельникова РАН (ФИРЭ РАН), Фрязино; старший научный сотрудник Института космических исследований РАН (ИКИ РАН), Москва.

Dmitry Mikhailovich ERMAKOV – Candidate of physical and mathematical sciences, leading scientist of the Fryazino branch of the Kotel'nikov Institute of Radio Engineering and Electronics of the Russian Academy of Sciences (FIRE RAS), Fryazino, Russia; senior researcher in the Space Research Institute of (IKI RAS), Moscow, Russia.

Current scientific interests: satellite microwave remote sensing, dynamics and energy of atmospheric mesoscale, synoptic and climatically significant processes, distributed data processing.

e-mail: dima@ire.rssi.ru



ЧЕРНУШИЧ Андрей Петрович – кандидат физико-математических наук, старший научный сотрудник Фрязинского филиала Института радиотехники и электроники им. В.А. Котельникова РАН (ФИРЭ РАН), Фрязино.

Andrey Petrovich CHERNUSHICH – Candidate of physical and mathematical sciences, senior researcher of the Fryazino branch of the Kotel'nikov Institute of Radio Engineering and Electronics of RAS (FIRE RAS), Fryazino, Russia.

Current scientific interests: satellite microwave remote sensing, distributed data processing, computer vision, image processing.

e-mail: andrey@fireras.su

Материал поступил в редакцию 16 декабря 2016 года