

УПРАВЛЕНИЕ КОЛЛЕКЦИЕЙ АСТРОМЕТРИЧЕСКИХ ПУЛЬСАРНЫХ ДАННЫХ В ОБЪЕКТНОЙ СРЕДЕ

А.Е. Авраменко

Пушчинская радиоастрономическая обсерватория, 142290, Пушино

Московской обл., ПРАО ФИАН

avr@prao.psn.ru

Для извлечения и отображения пространственно-временных свойств пульсаров по данным их многолетних астрометрических наблюдений была разработана компонентная объектная модель взаимодействия распределенных ресурсов коллекции, основанная на управлении данными и методами в неоднородной среде непосредственно из приложения. Решаются задачи трансформации и редуцирования измерительных данных различного типа с применением объектных методов приложения и процедур специализированных серверов. Рассматриваются вопросы связывания данных и сервисов на основе декомпозиции функций приложения и синтеза методов управления данными в предметной области для выявления исследуемых свойств. Приводится пример сценария обработки измерительных данных, инкапсулируемых через объектные методы приложения.

OBJECT ORIENTED ENVIROMENT FOR PULSAR ASTROMETRY DATA MANAGEMENT

A.E. Avramenko

Pushchino Radio Astronomy Observatory, Pushchino, Moscow region, 142290, Russia

avr@prao.psn.ru

It was designed the component object model of distributed resources interaction, based on management data and methods in heterogeneous ambience from application directly, for retrieving and presentation spatio-temporal properties of pulsars, extracted from astrometry observations data. There are decided the problems of transformation and reduction of different type observed data by using the object methods of application and special-purpose procedures of problem-oriented servers as well. There are considered the particularities of data and services consolidation, based on decomposition of application domain functions and synthesis of data management methods for retrieving the pulsar properties. The example of scenario of processing measured data, encapsulated through object methods of application, is given.

1. Введение

Уникальная коллекция астрометрических пульсарных данных ПРАО ФИАН содержит результаты многолетних наблюдений на крупнейших радиотелескопах России: БСА ФИАН (Пушино) и ТНА-1500 ОКБ МЭИ (Калязин). С целью ее интеграции с информационными ресурсами отечественных и зарубежных обсерваторий и центров были решены вопросы комплексирования информационных компонентов, независимо от источника, места размещения, используемых платформ [1]. На основе объектной модели коллекции были определены спецификации связывания и взаимодействия данных и сервисов (программ) непосредственно из клиентского приложения пользователя, обеспечивающие совместимость компонентов коллекции в процессе обработки данных и использования результатов совместно с Центром астрономических данных ИНАСАН (Россия), Институтом радиоастрономии М.Планка (MPIFR, Германия), Исследовательским центром коммуникаций (CRL, Япония).

Для отображения моментов прихода импульсов пульсаров на больших интервалах времени был разработан полный набор электронных форматов данных, поддерживающий атрибуты времени в последовательностях пульсарных событий на стадиях их регистрации, трансформации наблюдательных данных, моделирования частотно-временных свойств пульсаров в условиях распространения сигналов в межзвездной среде [2].

Дальнейшее развитие функциональных возможностей коллекции связано с извлечением существенных свойств предметной области на основе полученных наблюдательных и измерительных данных, консолидируемых через релевантные сервисы непосредственно в приложении. При этом существенно, что и консолидация ресурсов, и формирование запросов осуществляется в общем контексте проблемно-ориентированного приложения, определяющего содержание анализируемой предметной области [3]. На основе глубокого редуцирования и многопараметрической оптимизации больших объемов первичных («сырых») данных осуществляется переход от множества разрозненных, не проявляющих существенных закономерностей данных к статистически устойчивым модельным представлениям, обеспечивающим надежные оценки выявляемых свойств. Важно в среде пользовательского приложения обеспечить единство возможностей как формирования запросов, так и получения требуемых сведений, их оценки, визуализации представления извлекаемых свойств предметной области в виде гистограмм, сравнительных характеристик, закономерностей, оценок достоверности получаемых результатов, причем в течение относительно короткого, близкого к реальному масштабу, времени [4].

Реализация этих качеств ведет к существенному расширению функциональных возможностей коллекции, доступных ее пользователям. Применительно к интересующей нас области пульсарных исследований

можно сформулировать следующие основные направления решений, которые должны стать основой достижения рассматриваемых новых свойств коллекции:

- декомпозиция предметной области и синтез методов управления данными и сервисами в компонентной среде,
- интеграция методов и данных в процессе создания, поддержания коллекции, ее использования в приложениях,
- программирование сценариев извлечения свойств предметной области.

В работе рассмотрены особенности подхода к реализации этих методов, приводится пример практического решения.

2. Декомпозиция и синтез методов управления данными и сервисами

Объектно-ориентированная компонентная архитектура, поддерживающая взаимодействие данных и методов, обладает важным свойством естественной декомпозиции предметной области, что в сочетании с инкапсуляцией данных через релевантные методы позволяет моделировать поведение предметной области [5]. Развиваемая на основе объектного подхода концепция промежуточного слоя (middleware), в котором сосредоточены службы управления данными и методами (сервисами), обеспечивает возможность повторного интероперабельного использования компонентов в условиях информационной и реализационной неоднородности [6]. Набор используемых в коллекции типов данных соответствует содержанию предметной области разрабатываемых приложений, а синтезируемые методы управления данными обеспечивают выявление исследуемых свойств. При этом именно синтез методов требует основного внимания, поскольку определение методов представляет собой более сложный и трудоемкий процесс, чем описание типов данных, управляемых этими методами.

На Рис. 1 приведена композиционная структура управления пульсарными данными на технологических уровнях реализации приложений. В процессе наблюдений (Н) пульсаров в реальном масштабе времени осуществляется накопление данных (ДН), которые содержат исходную измерительную информацию о пульсарах (профили, моменты прихода импульсов). Полученные в наблюдениях данные затем трансформируются (ДТ) с целью выделения в явном виде измеряемых параметров с привязкой к соответствующим пространственно-временным системам отсчета и представления их в принятых унифицированных форматах для хранения, обмена и последующей обработки. Эти данные служат основой для моделирования (ДМ) физических свойств и поведения пульсаров в пространстве и времени. Все данные, как наблюдательные, так и результаты обработки, хранятся в базах данных (БД) без ограничения сроков, наряду с аналогичными внешними данными (ДВ), полученными другими обсерваториями в

совместных или аналогичных наблюдениях. Представление и интерпретация результатов (ДП) обеспечивает отображение свойств данных и особенности поведения моделей (СП).

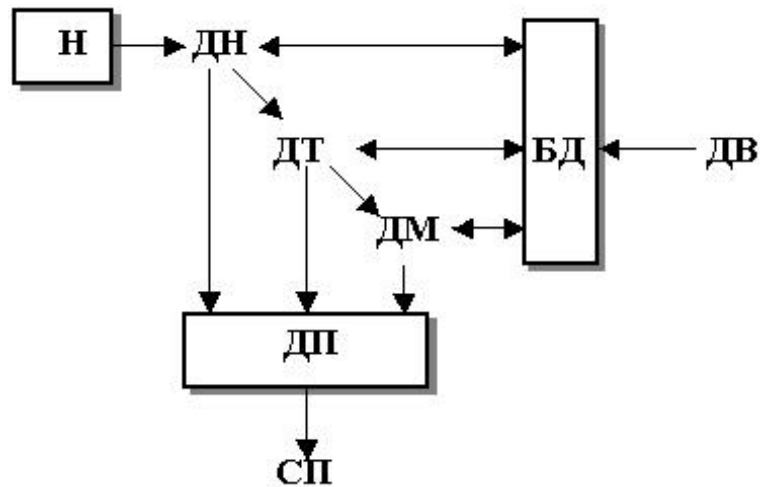


Рис. 1. Композиционная схема управления пульсарными данными

На Рис. 2 приводится обобщенная схема взаимодействия пульсарных данных и сервисов. Наиболее сложные, специализированные сервисы регистрации и трансформации (редуцирования) данных реализованы с использованием агентных (процедурных) программных средств, соответствующих многопараметрическим алгоритмам преобразования первичных наблюдательных данных к инвариантным значениям в принятых пространственно-временных системах отсчета. К ним относятся, в частности, программные пакеты PHAS для регистрации наблюдательных пульсарных событий в реальном времени, TIMAPR для определения параметров пульсаров, TIMSCALE для регулярных сличений локальной шкалы времени на радиотелескопе с опорной шкалой эталона. Эти программы, выполняющие фиксированные процедуры над входными данными, формируют исходные массивы, которые составляют содержательную основу баз пульсарных данных, отображаемую их атрибутами.

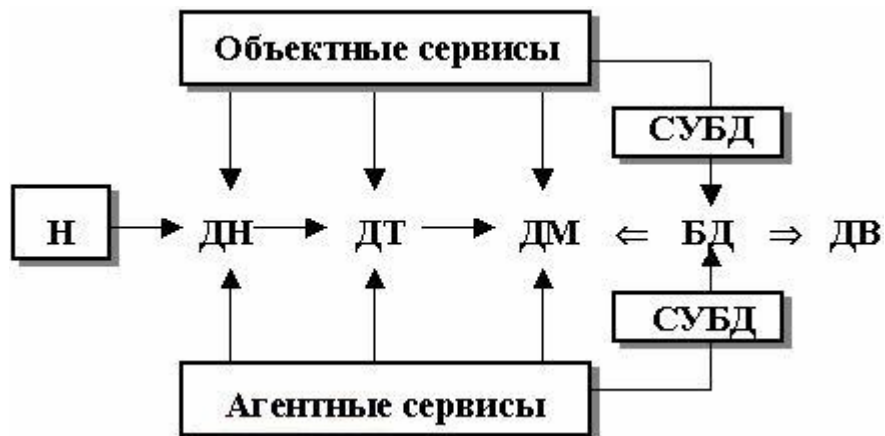


Рис. 2. Обобщенная схема взаимодействия пульсарных данных и сервисов

Вместе с тем, многочисленные сервисы редактирования, систематизации и представления данных, которые выявляют их семантическую значимость, функционируют непосредственно в объектной среде проблемно-ориентированного приложения, где осуществляется консолидация данных и сервисов, необходимых для отображения свойств и поведения предметной области. Данные и сервисы исходно либо локализованы в объектной среде приложения, либо распределены в локальном и/или глобальном пространстве.

Средства управления базами данных (СУБД) занимают промежуточное положение. Базы данных, как основная форма ведения наблюдательных, редуцированных данных и результатов обработки, распределены в локальном и/или глобальном пространстве по месту формирования и хранения, включая специализированные локальные серверы и серверы каталогов общего доступа. В каждом случае СУБД реализуется с учетом особенностей использования БД в приложениях и принятых на серверах БД схемах управления данными. Так, если данные подвергаются содержательному анализу средствами приложения, то соответствующие объектные сервисы приложения удобно использовать и для управления данными в БД (извлечение, редактирование, присоединение). Выборка данных на серверах БД общего назначения, непосредственно не относящихся к сфере управления объектных сервисов приложения, осуществляется с использованием провайдерских средств этих серверов.

3. Интеграция методов и данных в приложении

Рассмотренные особенности объектного подхода позволяют поддерживать в единой информационной среде как создание, так и выполнение проблемно-ориентированных приложений за счет интеграции данных и методов управления и обработки, с помощью которой достигается синтез модели предметной области с целью выявления ее свойств и поведения. Объектные технологии в сочетании с компонентной архитектурой позволяют интегрировать в приложении как концептуальную, так и реализационную схемы многосвязной предметной области. Объектный формализм (типы определений, наследование, методы) сопряжен непосредственно с понятиями и определениями предметной области, на основе которых реализуется единая инструментальная и технологическая среда приложения [7].

Консолидация (соединение) данных и сервисов осуществляется в приложении пользователя, моделирующем свойства и поведение предметной области. Данные и сервисы могут находиться как в приложении, так и, чаще всего, быть распределены в локальном и/или глобальном сетевом пространстве. Особенности условий консолидации данных и сервисов:

1. Данные, независимо от места их размещения, доступны приложению по локальной или глобальной сети.
2. Удаленные сервисы могут быть либо вызваны для выполнения в приложении, либо выполняться на месте непосредственно с данными.
3. Представление свойств данных в приложении осуществляется объектными сервисами-методами, которые инкапсулируют данные.
4. Управление в приложении инкапсулируемыми данными, размещенными в локальном и глобальном пространстве, производится как процедурными средствами удаленных серверов, так и объектными средствами приложения.

Приложение пользователя, отображающее результаты моделирования предметной области на основе данных коллекции, функционирует в оболочке клиента, которая в свою очередь должна располагать средствами доступа к данным и сервисам, с помощью клиент-серверных протоколов под управлением пользовательского интерфейса.

Объектно-ориентированная компонентная среда клиента выполняет функции промежуточного слоя (middleware) между приложением и операционной системой, обеспечивая взаимодействие пользователя с требуемыми ресурсами в процессе решения целевой задачи. Промежуточный слой в функционально завершенном виде представляет собой средства поддержки глобального пространства объектов на основе независимости спецификации их интерфейсов от конкретной реализации [6]. Такие средства реализованы, например, в компонентной архитектуре CORBA (Object Management Group) или DCOM (Microsoft), функционально аналогичной CORBA, но имеющей развитый объектно-ориентированный репозиторий компонентов, используемый в определяемых информационных моделях и поддерживаемый программируемым пользовательским интерфейсом в инструментальной среде Visual Basic приложения [8]. DCOM обладает свойством компонентной однотипности клиентской оболочки пользователя и операционной системы и поддерживает протоколы клиент-серверного взаимодействия в локальном и глобальном пространстве, содержащем неоднородные компоненты.

Из оболочки клиента пользователю доступны данные и удаленные сервисы и компоненты приложений, которые могут быть реализованы на серверах в рамках как объектной, так и агентной технологии. С помощью DCOM (Distributed COM) осуществляется удаленный доступ RPC (Remote Procedure Call) к объектным методам (сервисам) в локальном пространстве объектов. Компоненты объектных приложений доступны средствами пользовательского программируемого интерфейса API (Application Programming Interface). Доступ к агентным компонентам приложений и сервисам в глобальном пространстве осуществляется по протоколу HTTP с помощью API интернет-сервера (ISAPI) и CGI (Common Gateway Interface).

Протокол HTTP поддерживает управление доступом к данным на удаленном сервере в глобальном пространстве.

С помощью COM (DCOM) реализуется объектно-ориентированное управление данными с тремя основными типами серверов:

- сервер в самом процессе (in process): объекты реализуются динамически подключаемой библиотеке (DLL) и исполняются в том же процессе, что и клиент;
- локальный сервер: объекты реализованы в отдельном процессе, исполняющемся на той же машине, что и клиент;
- удаленный сервер: объекты реализованы в DLL либо в отдельном процессе, которые расположены на удаленном по отношению к клиенту компьютере.

Данные, находящиеся в локальном или глобальном пространстве имен, инкапсулируются в текущий процесс клиента методом *OpenFile* с указанием имени файла и пути к нему. В результате файл становится объектом данных, доступным для выполнения методов интерфейса *IdataObject*, включая *GetData* для передачи данных от объекта данных к его клиенту и *SetData* для передачи данных объекту данных от его клиента. Тем самым достигается инкапсуляция данных и методов в клиентском процессе с возможностью применения к данным требуемых для приложения пользователем объектных методов.

В частности, для инкапсуляции данных и методов объектно-ориентированного управления базами пульсарных данных исходная реляционная таблица, содержащая имена пульсаров, даты и результаты наблюдения, трансформируется, также, кстати, объектными методами приложения, в формат *.dbf* для БД, который поддерживается объектно-ориентированным методом извлечения данных с помощью SQL-запросов, формируемых непосредственно из приложения средствами графического программируемого интерфейса.

4. Программирование сценариев управления данными

Сценарии в рассматриваемой коллекции пульсарных данных представляют собой программы (макросы) обработки данных и отображения результатов, синтезированные их объектно-ориентированных сервисов, которые поддерживаются средствами визуального графического программирования и с помощью методов этих сервисов позволяют выполнять требуемую последовательность функциональных преобразований наблюдательных и измерительных данных. Это особенно удобно, когда исходные массивы данных однородны, структурно неизменны и многократно воспроизводятся в процессе наблюдений.

В качестве примера приводится сценарий обработки и представления результатов измерений хода местной шкалы времени на радиотелеско-

пах РТ-64 (Калязин) и БСА ФИАН (Пушино) по данным регулярных сличений с шкалой Госэталона времени и частоты России и спутниковой Global Positioning System (GPS). Сценарий объединяет сервисы, управляющие обработкой и отображением результатов сличений в виде, требуемом для использования этих результатов непосредственно в проблемно-ориентированных приложениях, в которых для прецизионной обработки наблюдательных данных хронометрирования пульсаров необходима точная привязка регистрируемых событий к временной шкале. Кроме того, содержащиеся в макросах методы позволяют получать количественные оценки метрологических параметров местных шкал времени на радиотелескопах для более точной их привязки интерполяционными методами.

Исходные данные сличений, формируемые на радиотелескопах в автоматическом режиме в виде последовательности измерительных выборок через определенные промежутки времени, инкапсулируются в объектно-ориентированную среду Visual Basic for Application, где с помощью внутренних сервисов средствами визуального графического программирования синтезируются требуемые проблемно-ориентированные методы трансформации, обработки и представления данных сличений. Временные ряды сличений, вместе с поправками Госстандарта России, которые отнесены к измеряемым моментам времени, и атрибутами времени (дата, часы, минуты) формируют БД сличений и результирующих поправок Службы Единого Времени (СЕВ) обсерватории. Путем последовательной сортировки данных по дате и времени выделяется на месячном интервале временной ряд поправок на один и тот же час каждой даты для вычисления суточного хода шкалы – основной метрологической характеристики СЕВ. Суточный ход представлен как в виде числового ряда, так и графически. Также формируется сводная таблица и график поправок во всех измеренных точках на оси времени, по которым берется поправка шкалы при обработке наблюдательных данных либо непосредственно в ближайшей точке, либо на основе интерполяции поправок хода шкалы. Ниже приведены основные компоненты Макроса 1, реализующего программу, а на Рис. 3 приведены графические результаты сличений.

Sub Макрос1()

' Сличение локальной шкалы времени

Range("E10:E69").Select

Selection.FormulaArray = "=-RC[-2]:R[59]C[-2]"

' Вычисление поправок шкалы СЕВ

Range("J81:J109").Select

Selection.FormulaArray = "=-RC[-5]:R[28]C[-5]-R[-1]C[-5]:R[27]C[-5]"

' Вычисление суточного хода шкалы СЕВ

Range("L89,A80:A109,J80:J109").Select

Range("J80").Activate

' Выборка данных для графического отображения

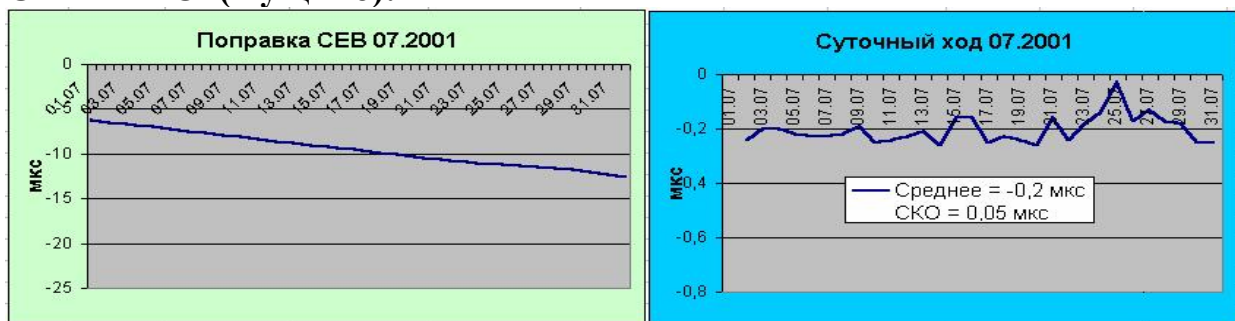
Charts.Add


```

ActiveChart.ChartType = xlLine
ActiveChart.SetSourceData
Source:=Sheets("Jul'01").Range("A80:A109,J80:J109") _
, PlotBy:=xlColumns
With ActiveChart
.HasTitle = True
.ChartTitle.Characters.Text = "Суточный ход 07.2001"
.Axes(xlValue, xlPrimary).AxisTitle.Characters.Text = "мкс"
End With
' Установка параметров графического отображения
ActiveChart.HasLegend = True
ActiveChart.Legend.Select
End Sub

```

СЕВ ПРАО (Пушино):



а) поправки местной шкалы

б) вариации хода местной шкалы

Рис. 3. Поправки местной шкалы и ее вариации на интервале 1 месяца

Оцениваются статистические характеристики шкалы СЕВ (Макрос 2):

- среднее значение хода шкалы $\bar{\tau}$ на месячном интервале,
- несмещенное стандартное отклонение σ_x шкалы СЕВ.

Результаты приведены в Таблице 1.

Sub Макрос2()

```

' Статистические характеристики локальной шкалы времени СЕВ
Range("K80").Select
Selection.FormulaArray = "=SUM(R[1]C[-1]:R[29]C[-1])/29"
' Вычисление среднего значения хода шкалы СЕВ
Range("L81:L109").Select
Selection.FormulaArray = "=RC[-2]:R[28]C[-2]-R[-1]C[-1]"
' Вычисление текущего отклонения шкалы СЕВ от среднего
Range("M80").Select
ActiveCell.FormulaR1C1 = "=SQRT(DEVSQ(R[1]C[-1]:R[29]C[-1])/28)"
Range("M81").Select
' Вычисление несмещенного стандартного отклонения шкалы СЕВ
End Sub

```

Таблица 1

Дата	Время	СЕВ-GPS	Поправка GPS - UTC(SU)	Поправка СЕВ [UTC(SU)]	Суточный ход	$\bar{\tau}$	$\tau_i - \bar{\tau}$	σ_x
		Δ_1 , мкс	Δ_2 , мкс	Δ_3 , мкс	τ , мкс			
01.07	8.00	6,26		-6,26		-0,2040		0,0500
02.07	8.00	6,5		-6,5	-0,24		-0,0360	
03.07	8.00	6,7		-6,7	-0,2		0,0040	
04.07	8.00	6,9		-6,9	-0,2		0,0040	
05.07	8.00	7,12		-7,12	-0,22		-0,0160	
06.07	8.00	7,35		-7,35	-0,23		-0,0260	
07.07	8.00	7,58		-7,58	-0,23		-0,0260	
08.07	8.00	7,8		-7,8	-0,22		-0,0160	
09.07	8.00	7,99		-7,99	-0,19		0,0140	
10.07	8.00	8,24		-8,24	-0,25		-0,0460	

Полученные на основе исходных данных таблицы и графики группируются с разбивкой на месяцы в пределах года, и в виде Бюллетеня СЕВ доступны на Web-сервере [9].

5. Заключение

Рассматриваемый подход к управлению коллекцией астрометрических пульсарных данных с целью извлечения и отображения пространственно-временных свойств пульсаров основан на связывании (консолидации) компонентов – данных и методов – в объектной среде проблемно-ориентированного приложения. Данные, независимо от используемых платформ, средств формирования и места расположения серверов, специфицируются по типам и форматам и инкапсулируются объектными методами, которые синтезируют в приложении функции управления данными, выявления и моделирования свойств, отображения результатов. Объектная среда приложения, обладая компонентной однотипностью клиентской оболочки пользователя и операционной системы, обеспечивает технологическое единство создания и выполнения проблемно-ориентированных приложений на основе графического программируемого интерфейса и унифицированных протоколов клиент-серверного взаимодействия в локальном и глобальном пространстве.

Разработанные сценарии управления данными в объектной среде позволяют реализовать требуемую последовательность функциональных преобразований наблюдательных и измерительных данных и представление результатов предметной области.

Выполненная работа будет положена в основу следующего этапа развития интегрированной коллекции пульсарных данных, предусматривающего систематизацию новых знаний о физических свойствах пульсаров, извлекаемых из наблюдательных данных в среде пользовательского приложения, оценки достоверности, сравнительных характеристик, гистограмм, закономерностей, визуализацию представления получаемых результатов, причем в течение относительно короткого, близкого к реальному масштабу, времени.

Работа поддерживается грантом РФФИ № 00-07-90033.

Литература

1. А.Е.Авраменко, О.Б.Длужневская, О.В.Дорошенко, В.А.Потапов, А.Е.Родин. Объектная модель коллекции астрометрических пульсарных данных. // Труды Второй Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». – Протвино, сентябрь 2000, с.193-197.
2. А.Е.Авраменко, О.В.Дорошенко, Ю.П.Илясов, В.А.Потапов, О.Б.Длужневская. Событийная непрерывность коллекции астрометрических пульсарных данных. // Сборник трудов Третьей Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». – Петрозаводск, сентябрь 2001, с.192-196.
3. T.Palpanas. Knowledge Discovery in Data Warehouses. SIGMOD Record, v.29, No.3, Sept.2000, <http://www.acm.org/sigmod/record/issues/0009/>
4. index.html.
5. P.J.Naas, J.M.Hellerstein. Online Query Processing. SIGMOD 2001 Tutorial, <http://control.cs.berkeley.edu/sigmod01/>.
6. В.П.Иванников, К.В.Дышлевой, С.Г.Манжелей, Л.Б.Соловская, А.Б.Шебуняев. Распределенные объектно-ориентированные системы. // Труды Института системного программирования РАН. – том 1. М., 2000.
7. Д.О.Брюхов, В.И.Задорожный, Л.А.Калиниченко, М.Ю.Курошев, С.С.Шумилов. Интероперабельные информационные системы: архитектура и технологии. – СУБД, 4, 1995.
8. A.Rosenthal. Where Will Object Technology Drive Data Administration? SIGMOD Record, v.27, No.1, March 1998. <http://www.acm.org/sigmod/record/issues/9803/index.html>.
9. P.A.Bernstein et al. The Microsoft Repository. Very Large Database Conf., Athens, Greece, 1997, <http://www.acm.org/sigmod/dblp/db/conf/vldb/BernsteinHSSZ97.html>
10. Bulletin of Local Time Service PRAO, <http://psun32.prao.psn.ru/data/TS/index.html>