

В настоящее время существует множество технологий, позволяющих создавать web-приложения для доступа к удаленным базам данных. Однако современная технология ASP.NET отличается от них высокой степенью интеграции с серверными продуктами, а также с инструментами Microsoft для разработки, доступа к данным и обеспечения безопасности [4]. Именно эта технология принята в качестве основного инструмента для реализации web-приложений архива проектных решений. При этом для хранения проектных данных, составляющих основу АПР, предполагается использовать платформу SQL Server 2008, предназначенную для интеграции корпоративных данных, отличающуюся хорошей масштабируемостью, отличными показателями извлечения, преобразования и загрузки, широкими возможностями интеграции. При ее использовании упрощается управление данными, собранными из самых различных источников.

Таким образом, предложенный подход к организации АПР САПР позволяет автоматизировать процедуру сохранения, поиска и повторного использования проектных решений схемотехнических САПР.

СПИСОК ЛИТЕРАТУРЫ

1. Разевиг В. Д. Система сквозного проектирования электронных устройств DesignCenter 8.0. М.: Солон, 1999.
2. Разевиг В. Д. Система схемотехнического моделирования Micro-Cap V. М.: Солон, 1997.
3. Болотовский Ю. Б., Таназлы Г. И. OrCAD. Моделирование "Поваренная" книга. М.: Солон-Пресс, 2005.
4. Мак-Дональд М., Фримен А., Шпушта М. Microsoft ASP.NET 4.0 с примерами на C# 2010 для профессионалов / Пер. с англ. М.: Вильямс, 2011.

G. D. Dmitrevich, A. I. Laristov, Al-Shameri Yazid Mohammed Abdulrahman

MODEL DATA ARCHIVE PROJECT MAKING SCHEMATIC CAD

The questions of the organization web-based archive of circuit design solutions CAD. Discusses the reusability of design solutions derived in new development and modernization of the corresponding node or block CEA. Analyzed the contents of the archive project decisions and propose a model of data backup in the form of ER-diagrams provide a framework for future implementation archive in a relational database environment.

Schematic CAD, archive design solutions, the data model, ER-diagram

УДК 519.863

В. С. Абатуров

АРХИТЕКТУРА ПРОМЫШЛЕННОЙ АНАЛИТИЧЕСКОЙ ПОДСИСТЕМЫ ИЗВЛЕЧЕНИЯ ЗНАНИЙ

Предлагается вариант архитектурного решения аналитической подсистемы для промышленных и бортовых применений на основе СУБД PostgreSQL. Классические этапы Data Mining, включающие обучающую фазу, фазу тестирования и фазу применения, предлагается реализовать с помощью хранимых процедур СУБД PostgreSQL. Архитектура подсистемы подчинена требованиям SQL/MM и PMML.

Аналитическая подсистема, извлечение знаний, аналитическая база данных

Современные системы управления технологическими процессами требуют обработки и анализа больших объемов информации. Данная проблема давно стала критической не только в областях, связанных с аналитической обработкой данных (искусственный интел-

лект, системы поддержки принятия решений, техническое зрение, мультимедиа-технологии и др.), но и в естественно-научных дисциплинах. Например, база данных снимков для астрономии давно превзошла размеры нескольких петабайт (1 петабайт = 1024 терабайт) [1].

В настоящее время технология Data Mining обладает наиболее развитыми методами извлечения знаний из «сырых данных». В широком смысле под термином Data Mining понимается совокупность методов для обнаружения ранее неизвестных, нетривиальных, практически полезных и доступных для интерпретации знаний в необработанных данных («сырых данных»). При этом предполагается, что исходная база имеет достаточно большой объем данных и содержит скрытые знания. Процесс извлечения знаний представлен на рис. 1.

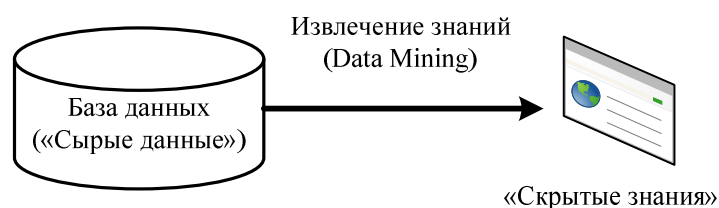


Рис. 1

Реализация технологии Data Mining для промышленных и бортовых систем связана с решением ряда принципиальных вопросов, к которым относятся: выбор архитектуры, системных интерфейсов, обеспечение сервисных возможностей, безопасности, надежности и высокого быстродействия.

Тенденция последних лет в развитии этого направления заключается в интеграции средств аналитической обработки, выделения знаний, управления метаданными и визуализации результатов на одной программной аналитической платформе.

В настоящее время рынок аналитических систем экспоненциально развивается. В этом развитии принимают участие практически все крупнейшие корпорации¹. На этом рынке представлены такие крупные зарубежные компании, как IBM Cognos, MicroStrategy, Oracle, SAS, Microsoft, а также российские разработчики: фирмы BaseGroup Labs с пакетом Deductor, Intersoft Lab с платформой хранилищ данных и аналитической платформой «Контур», фирма «Прогноз» с аналитической платформой PROGNOZ Platform.

Архитектура аналитической подсистемы. В настоящей статье предлагается вариант архитектурного решения аналитической подсистемы для промышленных и бортовых применений на основе СУБД PostgreSQL. На рис. 2 представлена архитектура аналитической подсистемы извлечения знаний.

PostgreSQL – объектно-реляционная система управления базами данных с открытыми кодами. Данная СУБД позволяет применять пользовательские хранимые процедуры для выполнения процессов извлечения знаний. Кроме того данная СУБД позволяет создавать пользовательские типы данных, необходимые для хранения и взаимодействия функциональных элементов, необходимых для организации процесса извлечения знаний.

¹ <http://www.kdnuggets.com> – Copyright © 2013 KDnuggets.

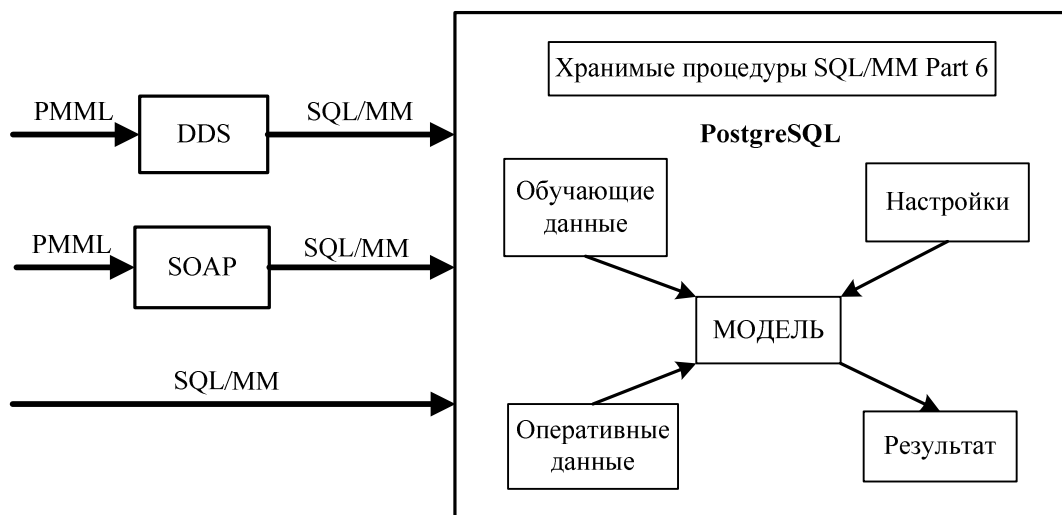


Рис. 2

Взаимодействие с приложениями для аналитической подсистемы данной архитектуры осуществляется с помощью трех интерфейсов:

- SQL/MM¹ – расширение языка SQL для управления процессами извлечения знаний;
- SOAP Version 1.2² – протокол коммуникаций (SOAP) между интернет-приложениями, основан на XML и HTTP;
- DDS – открытый стандарт распределенного сервиса для систем реального времени (OMG)³.

Ключевую роль в данной архитектуре играет язык PMML⁴ – стандарт на основе XML по представлению и передаче моделей Data Mining (PMML). Данный стандарт применяется не только для представления и хранения всех моделей знаний, но и для взаимодействия с аналитической подсистемой.

Предложенная архитектура позволяет выполнять классические этапы Data Mining, включающие обучающую фазу, фазу тестирования и фазу применения с помощью хранимых процедур СУБД PostgreSQL, выполненных в соответствии со стандартом SQL/MM. На рис. 3 представлена обучающая фаза, на рис. 4 – фаза тестирования, на рис. 5 – фаза применения.

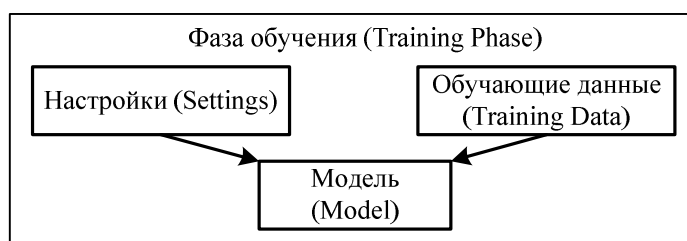


Рис. 3

¹ ISO/IEC 13249-6-2006, SQL/MM Part 6.

² SOAP Version 1.2, 2007, Рекомендация W3C.

³ OMG Document formal/07-01-01, Version 1.3.

⁴ PMML Version 4.1, 2012, Data Mining Group (DMG) <http://www.dmg.org/>.

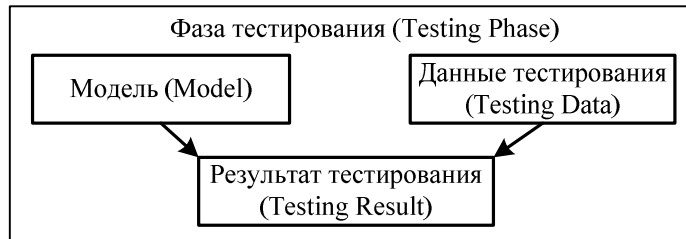


Рис. 4

Фаза обучения – этап, на котором строится вычислительная модель интеллектуального анализа данных. Испытательная фаза – этап, на котором осуществляется проверка качества предсказания на основе построенной модели. Прикладная фаза – этап, на котором строка оперативных данных оценивается на основе модели, построенной на этапе обучения.

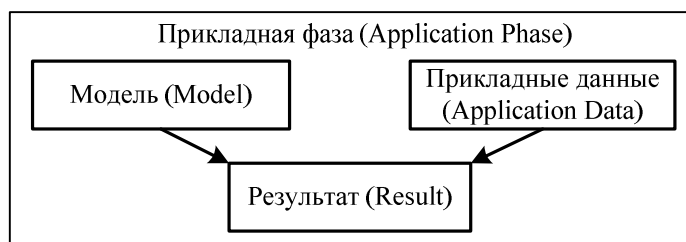


Рис. 5

Из рис. 5 видно, что в данной архитектуре СУБД не только организует выполнение процессов извлечения знаний, но также выполняет функцию хранения обучающих данных, прикладных данных, данных тестирования, настроек, моделей знаний и других функциональных элементов.

Использование данной аналитической подсистемы предполагает наличие системного аналитика и системного администратора. На рис. 6 представлена схема сценариев использования аналитической подсистемы.

Системный аналитик осуществляет настройку процессов извлечения знаний, инициирует построение моделей знаний. В его функции входит: определение алгоритмов, которые необходимо использовать, их настройки, выбор данных для обучения и тестирования, оценка качества полученных моделей.

Системный администратор производит инсталляцию аналитической подсистемы, выполняет администрирование правами доступа пользователей, управляет базой моделей (удаление устаревших моделей, удаление неиспользуемых моделей).

Подсистема поддержки принятия решения применяет полученные модели для выполнения прикладной фазы процесса извлечения знаний. Полученный результат используется для выработки рекомендаций и решений в процессе управления.

Роль хранилища общих данных выполняет база данных под управлением PostgreSQL. Данная база содержит анализируемые данные: таблицы обучающих данных, таблицы тестовых данных, таблицы прикладных данных; объекты, необходимые для выполнения процессов анализа данных: таблицы настроек для различных алгоритмов, таблицы задач извлечения знаний; объекты, получаемые в результате анализа данных, т. е. модели знаний: таблицы моделей знаний.

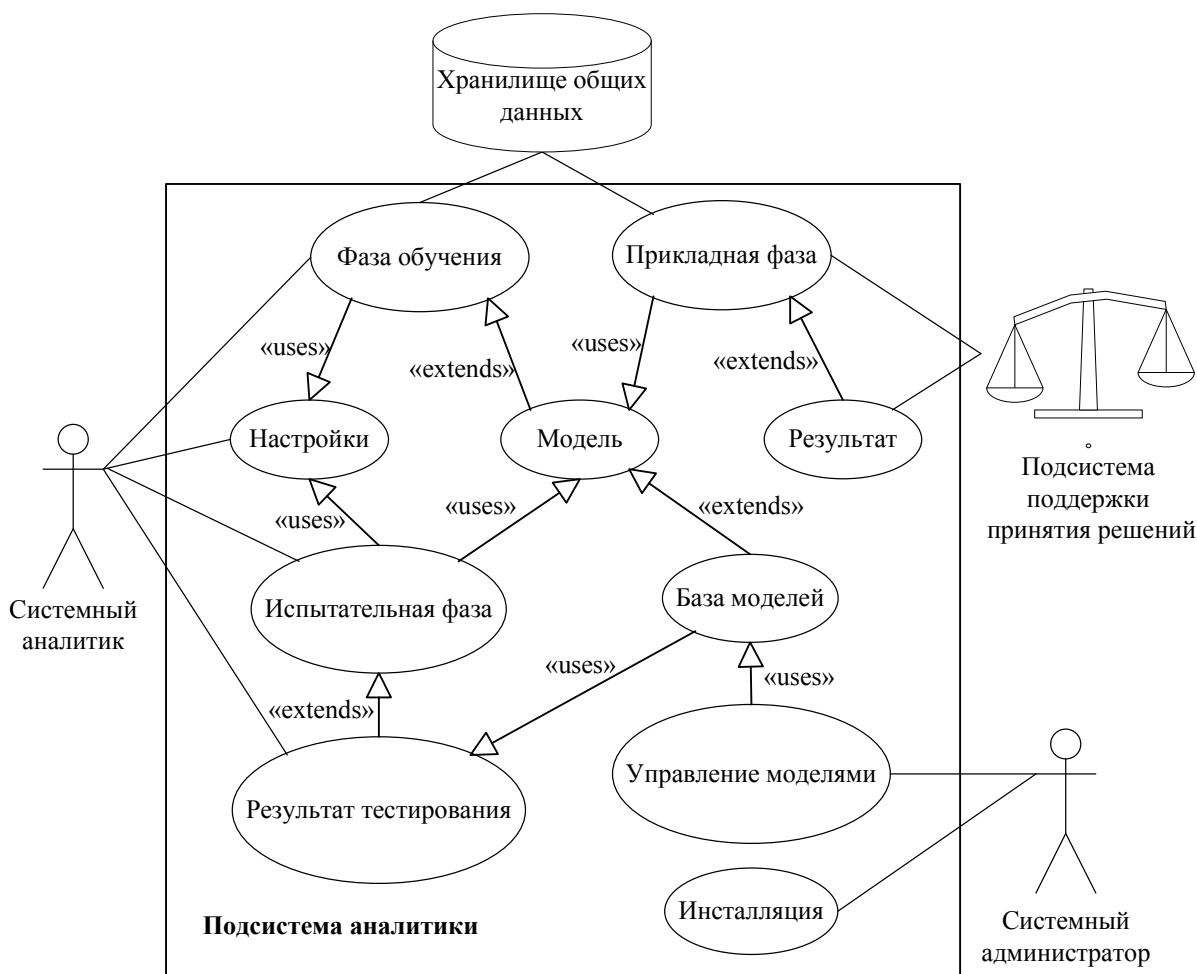


Рис. 6

В соответствии со стандартом SQL/MM различные модели знаний и вспомогательные объекты для их построения, ввиду своей структурной разнородности, должны храниться в различных объектах и таблицах в зависимости от алгоритма, порождающего модель. На рис. 7 представлена диаграмма классов базы данных, используемая для построения деревьев решений (Decision Trees).

Данная диаграмма классов затрагивает компоненты, необходимые для построения, хранения и использования только для алгоритма «Деревья решений». База данных содержит аналогичные компоненты для других алгоритмов. В таблице представлен перечень алгоритмов Data Mining, реализацию которых допускает данная архитектура. Данный перечень алгоритмов не является полным. Предложенная архитектура позволяет производить модификацию базовых алгоритмов, тем самым наращивая функционал аналитической подсистемы. Это достигается благодаря гибкости стандарта представления моделей знаний PMML.

Несмотря на то, что стандарт SQL/MM предусматривает наличие трех фаз в процессе извлечения знаний (фаза обучения, фаза тестирования и фаза применения), данный стандарт разрешает ограничиться одной фазой. Например, при построении когнитивной карты или статистического портрета требуется только фаза обучения.

Для выполнения процесса извлечения знаний используются следующие функциональные компоненты СУБД PostgreSQL: типы данных, хранимые процедуры, исключения. Данные компоненты позволяют формировать SQL/MM – скрипт для проведения анализа данных. На рис. 7 представлены функциональные компоненты SQL/MM – аналитической подсистемы.

Метод	Назначение
Ассоциативные правила	Прогнозирование событий
Кластеризация	Разбиение данных на классы
Деревья решений	Классификация состояний
Регрессия	Выявление скрытых зависимостей
Статистический портрет	Характеристики вероятностной модели
Метод k ближайших соседей	Классификация образов
Наивный байесовский классификатор	Классификация состояния на вероятностных моделях
Логические правила	Автоматическая экспертиза
Нейронные сети	Классификация состояний
Временные последовательности	Прогнозирование временных рядов
Текстовые модели	Семантический анализ
Машина опорных векторов	Классификация состояний
Сбалансированная система показателей	Планирование операционной деятельности и контроля их достижения

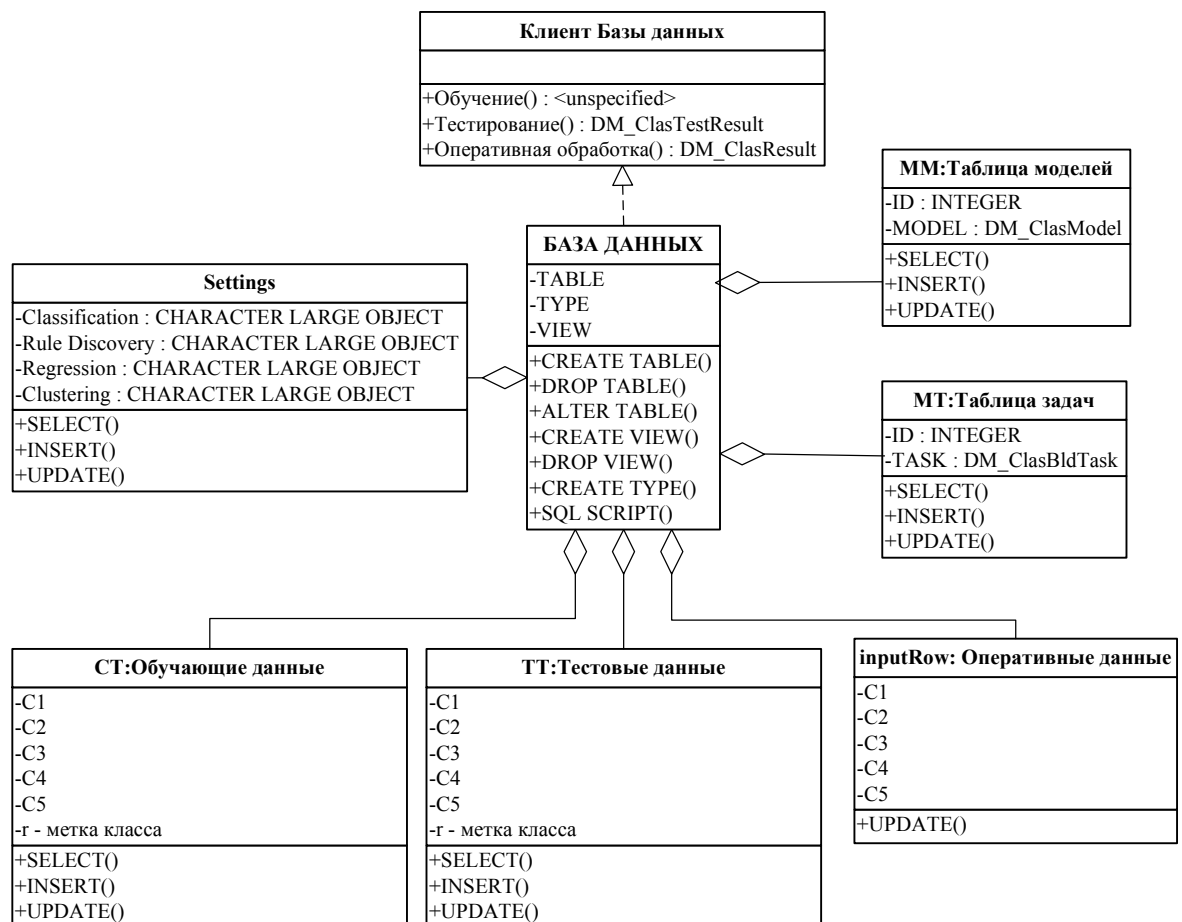


Рис. 7

Типы данных предназначены для хранения различной информации, необходимой при анализе данных. На рис. 8 в качестве примера представлены типы DM_MiningData и DM_LogicalDataSpec. Данные типы предназначены для представления метаданных анализируемых данных. Хранимые процедуры служат для выполнения взаимодействия между различными типами, а также для запуска алгоритмов анализа данных. Исключения типизированы и служат для представления пользователю информации об ошибках, допущенных при построении скриптов анализа данных, несоответствии алгоритмов и используемых моделей знаний и многих других ситуациях. Данные компоненты, реализованные с помощью средств PostgreSQL, позволяют формировать SQL/MM-скрипт, предназначенный для запуска процессов извлечения знаний. На рис. 8 представлен пример SQL/MM-скрипта для построения задачи классификации данных.

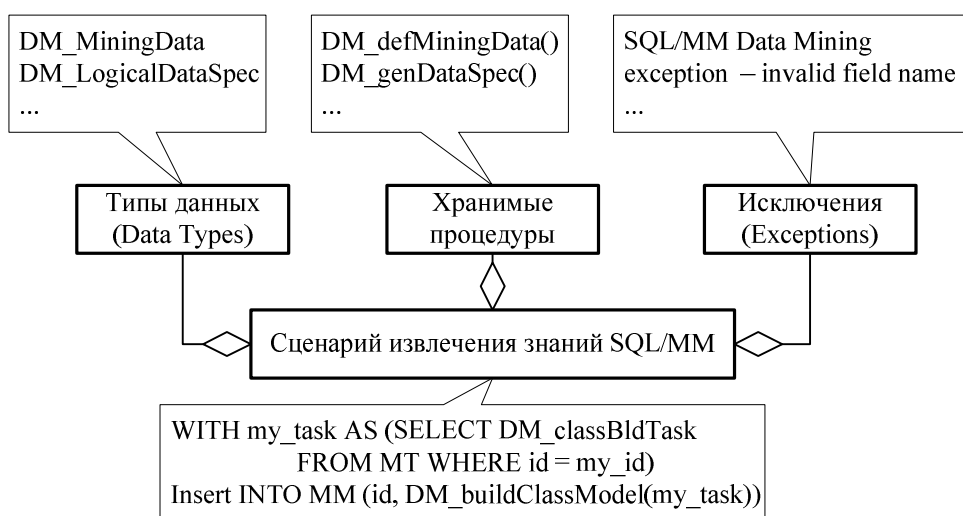


Рис. 8

Тактико-технические характеристики аналитической подсистемы. Поскольку в качестве базовой платформы была выбрана СУБД PostgreSQL, аналитическая подсистема, реализованная в соответствии с предложенной архитектурой, будет иметь следующие тактико-технические характеристики:

- обработка масштабных массивов разнородной информации (до 32 Тбайт);
- расширяемость и масштабируемость аналитики;
- многоплатформенность;
- контроль целостности данных;
- интерфейсы доступа к языкам программирования высокого уровня (С, С++, R, Java и другие языки программирования);
- методы доступа к данным JDBC, ODBC;
- триггеры и правила для управления процессами;
- система управления правами доступа и авторизации;
- шифрование трафика;
- аппаратное ускорение вычислений [2].

Преимуществами предложенной аналитической подсистемы является гибкость применения алгоритмов извлечения знаний, а также их расширяемость. Эти возможности достигаются во многом благодаря стандарту представления моделей знаний PMML и стандарту управления процессами извлечения знаний SQL/MM.

Предложенная архитектура аналитической подсистемы извлечения знаний, основанной на базе СУБД PostgreSQL, позволяет эффективно организовать процессы аналитики и извлечения знаний в различных предметных областях, таких, как искусственный интеллект, системы поддержки принятия решения, техническое зрение, мультимедиа-технологии и др. СУБД PostgreSQL применяется как для хранения анализируемых данных, так и для хранения извлеченных знаний. Для представления моделей в СУБД предлагается использовать стандарт PMML. Стандарт SQL/MM позволяет предоставлять унифицированный доступ пользователей к аналитической подсистеме, а также унифицированный API для различных программных средств.

СПИСОК ЛИТЕРАТУРЫ

1. Барсегян А. А. Анализ данных и процессов. 3-е изд. СПб.: БХВ-Петербург, 2009.
2. Jamieson S. Micro Telecommunications Computing Architecture. Short Form Specification – Micro Telecommunications Computing Architecture. PICMG® MTCA.0 R1.0, Sept. 21st. 2006.

V. S. Abaturov

THE ARCHITECTURE OF INDUSTRIAL ANALYTICAL DATA MINING SYSTEMS

The variant of the architectural solutions of analytical subsystems for industrial and aircraft applications based on PostgreSQL database is presented in the article. It is proposed to implement Data Mining classical phases, including training phase, testing phase and application phase using the stored procedures of PostgreSQL database. The architecture of the subsystem corresponds the requirements of the SQL/MM and PMML.

Analytical subsystem, Data Mining, analytical database

УДК 519.713; 530.145; 004.27; 004.414

И. В. Матвеева, В. А. Калмычков

ГИБКИЕ МАРШРУТЫ ПРОЕКТИРОВАНИЯ КВАНТОВЫХ ВЫЧИСЛИТЕЛЬНЫХ УСТРОЙСТВ В VI-СРЕДАХ САПР

Рассмотрено применение гибких маршрутов проектирования на примере построения многоуровневых спецификаций квантовой цепи для сумматора.

Кубит, квантовая цепь, сумматор, гибкий маршрут проектирования

Нанотехнологии все активнее становятся существенной составляющей нашей жизни, предлагая новые решения для традиционных задач. Важные изменения происходят и в области компьютерных технологий, при этом выход элементной базы на нанометрический уровень четко обозначил предельные возможности по дальнейшему ее развитию. По оценкам специалистов Российского квантового центра наиболее перспективны в качестве технологий реализации квантовых цепей технологии, основанные на атомах, ионах, сверхпроводящих схемах, одиночных спинах в твердотельных системах. Актуальной становится задача исследования особенностей компьютерной элементной базы именно на атомарном уровне ее реализации, а также развитие соответствующих методов проектирования. Необходим переход от реализации классических способов (алгоритмов) обработки данных к новым вариантам, учитывающим особенности процессов в принципиально иной технологической среде. Для описания таких подходов может быть использована модель представления квантовых цепей [1].

Предложения по системам автоматизации проектирования устройств в нотации квантовых цепей существуют, но имеют скорее архитектурный характер при наличии и предложений по их реализации. Например, Методы и др. [2] предложена однородная квантовая архитектура на основе логического массива и инструментарий для автоматической