

На правах рукописи



**Вейбер Вадим Викторович**

**Алгоритмическое и программное обеспечение  
инструментальной системы для интеграции  
производственных данных  
нефтегазодобывающей компании**

05.13.11 — Математическое и программное обеспечение  
вычислительных машин, комплексов и компьютерных  
сетей

**А в т о р е ф е р а т**  
**диссертации на соискание ученой степени**  
**кандидата технических наук**

Томск — 2011

Работа выполнена на кафедре вычислительной техники ФГБОУ ВПО  
«Национальный исследовательский Томский политехнический университет»

Научный руководитель: доктор технических наук, профессор,  
Заслуженный деятель науки РФ  
**Марков Николай Григорьевич**

Официальные оппоненты: доктор технических наук, профессор  
**Ехлаков Юрий Поликарпович**

кандидат технических наук  
**Напрюшкин Александр  
Алексеевич**

Ведущая организация: **Новосибирский государственный  
технический университет**

Защита состоится 22 декабря 2011 г. в 15 часов на заседании Совета по защите докторских и кандидатских диссертаций Д 212.269.06 при ФГБОУ ВПО «Национальный исследовательский Томский политехнический университет» по адресу: 634034, г. Томск, ул. Советская, 84/3, Институт кибернетики ТПУ, ауд. 214.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО «Национальный исследовательский Томский политехнический университет» по адресу: 634034, г. Томск, ул. Белинского, 55.

Автореферат разослан « 21 » ноября 2011 г.

Ученый секретарь Совета по защите докторских  
и кандидатских диссертаций Д 212.269.06  
кандидат технических наук, доцент



М.А. Сонькин

## Общая характеристика работы

**Актуальность темы.** В процессе развития крупной промышленной компании многократно возрастают ее информационные потоки и все более значимой становится правильная организация этих потоков внутри компании. Анализ проблем автоматизации нефтегазодобывающих компаний (НГДК) показывает, что «лоскутная автоматизация», т. е. наличие нескольких информационных систем (ИС) и специализированных комплексов программ разных поколений, присуща многим из них. Это объясняется тем, что никакая интегрированная ИС даже класса ERP (Enterprise Resource Planning – планирование ресурсов предприятия) не покрывает в перспективе всех потребностей в автоматизации такой компании, тем более ее специальных потребностей, связанных с особенностями производства непрерывного типа. Сегодня автоматизация НГДК, обеспечиваемая разнообразными эксплуатируемыми и вновь внедряемыми ИС и комплексами программ, как правило, часто сводится к автоматизации лишь отдельных участков и цехов производства, а не производственного цикла компании в целом. При этом многие производственные задачи решаются неэффективно или не решаются вовсе без совместного использования данных сразу нескольких ИС и (или) комплексов программ. На текущий момент уровень интеграции таких ИС и комплексов программ весьма невысок: взаимосвязанные бизнес-процессы компании автоматизированы с использованием многочисленных унаследованных и устаревших решений по интеграции разнородных ИС и комплексов программ. Использование таких интеграционных решений ведет к большим рискам в части обеспечения надежности ИТ-поддержки бизнеса НГДК.

Все это указывает на то, что в нефтегазовой отрасли научная проблематика по интеграции эксплуатируемых в НГДК ИС и комплексов программ и, в конечном итоге, создания единого информационного пространства (ЕИП) таких компаний, особенно актуальна.

Существуют различные подходы к интеграции ИС и комплексов программ компании: по данным, по функциям, по интерфейсам и т.д. Наиболее перспективным и часто используемым из них является подход к интеграции ИС и комплексов программ по данным. Среди ученых, изучающих проблемы интеграции данных в крупных компаниях, следует отметить Д.В. Торшина, И.С. Михайлова, М.Р. Когаловского, Л.А. Калиниченко, В.В. Липаева, Е.Н. Филинова, А.В. Горбатова, А. Naumenko, E. Rahm и других.

Анализ результатов работ этих исследователей и анализ ряда выполненных проектов интеграции ИС производственного назначения показал, что проблема интеграции производственных данных промышленных предприятий, в том числе компаний нефтегазовой отрасли, представляется недостаточно изученной, а большинство проектов по интеграции данных в крупных компаниях не завершились успешно из-за отсутствия общих подходов и способов решения этой проблемы. Более того, практически значимые алгоритмы и инструментальные программные средства для решения проблемы интегра-

ции производственных данных, учитывающие специфику нефтегазовой отрасли, сегодня на рынке программного обеспечения отсутствуют.

**Целью диссертационной работы** является создание алгоритмического и программного обеспечения инструментальной системы для интеграции производственных данных информационных систем и комплексов программ, производственного назначения, эксплуатируемых в нефтегазодобывающих компаниях.

Для реализации поставленной цели необходимо последовательное решение следующих задач.

1. Разработка принципов построения и архитектуры инструментальной системы для интеграции производственных данных (СИПД) НГДК, включая интеграцию технологических данных различных АСУТП, используемых в нефтегазовой отрасли.

2. Создание интеграционной модели производственных данных (ИМПД) НГДК на основе предложенных принципов построения и сформулированных требований к разрабатываемой инструментальной системе.

3. Разработка алгоритмического обеспечения инструментальной СИПД. Решение данной задачи предполагает также исследование эффективности предлагаемых алгоритмов.

4. Разработка программного обеспечения (ПО) инструментальной СИПД. Результатом решения этой задачи должны явиться программные средства, созданные с учетом разработанных принципов и архитектуры инструментальной системы и реализующие предложенные алгоритмы.

5. Аprobация и внедрение разработанной инструментальной системы при решении практических задач создания конкретных СИПД и интеграции с их помощью производственных данных современных НГДК.

**Методы исследований.** В работе использованы методы теории множеств, объектно-ориентированного проектирования ПО и проектирования БД, методология структурного анализа и проектирования SADT.

**Научную новизну полученных в работе результатов определяют:**

1. Архитектура инструментальной СИПД, отличительной особенностью которой является использование подхода к интеграции данных на основе интеграционной модели производственных данных и принципов сервисно-ориентированной архитектуры (англ. Service Oriented Architecture, SOA).

2. Интеграционная модель производственных данных НГДК, созданная на основе отраслевого стандарта PRODML и подхода EAV («Сущность-Атрибут-Значение») и позволяющая гибко описывать объекты производства.

3. Алгоритм сопоставления схем данных участвующих в интеграции производственных ИС и комплексов программ со схемой данных ИМПД компании, дающий более высокое качество сопоставления схем по сравнению с рядом аналогичных алгоритмов широко известной программной среды СОМА+ за счет оригинального структурного сопоставления схем данных.

4. Способ сбора и передачи (доставки) на верхние уровни управления компанией технологических данных различных используемых АСУТП, от-

личительной особенностью которого является обеспечение высокой производительности по сравнению с аналогами и обеспечение требуемого уровня информационной безопасности в современной НГДК.

**Практическая ценность и реализация результатов работы.** Практически значимыми являются созданные ИМПД, алгоритмы и программные средства инструментальной СИПД для современных НГДК. Объем исходного кода разработанных программных средств составляет более 17 000 строк на языках C#, Delphi, T-SQL и PL/SQL. Часть программ имеет государственную регистрацию.

Созданные ИМПД, алгоритмы и программные средства инструментальной СИПД были использованы при выполнении х/д 8-02/08 и х/д 8-05/09 между Институтом «Кибернетический центр» Томского политехнического университета и ОАО «Томскгазпром» и внедрены в виде конкретной СИПД, использованной в ОАО «Томскгазпром» при решении ряда задач интеграции ИС и АСУТП по данным. Интеграции подлежали корпоративная геоинформационная система управления производством (КГСУ) «Магистраль-Восток», ИС обработки и интерпретации геолого-геофизических данных «BASPRO», информационно-аналитическая система для моделирования промысловых трубопроводов «OISPipe» и ряд АСУТП промыслов ОАО «Томскгазпром». Результаты внедрения подтверждены соответствующими актами.

#### **Основные положения, выносимые на защиту:**

1. Предложенная интеграционная модель производственных данных позволяет гибко описывать в виде структур данных производственные объекты добывающих компаний нефтегазовой отрасли.

2. Разработанный алгоритм сопоставления схем данных ИС и комплексов программ обеспечивает более высокую точность сопоставления схем по сравнению с аналогичными алгоритмами широко известной программной среды СОМА+.

3. Предложенный способ сбора и передачи (доставки) технологических данных различных АСУТП нефтегазодобывающей компании на верхние уровни управления компанией, базирующийся на принципах сервисно-ориентированной архитектуры, обеспечивает более высокую производительность при доставке данных по протоколу OPC по сравнению с аналогами.

4. Разработанная архитектура, созданное алгоритмическое и программное обеспечение инструментальной СИПД позволяют создавать конкретные СИПД и с их помощью эффективно решать интеграционные задачи по производственным данным в добывающих компаниях нефтегазовой отрасли.

**Апробация работы.** Основные результаты работы докладывались и обсуждались на VI и VII Всероссийской научно-практической конференции студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии» (г. Томск, 2008 г.; г. Томск, 2009 г.), XV Международной научно-практической конференции студентов, аспирантов и молодых учёных «Современная техника и технологии» (г. Томск, 2009 г.), VII и VIII

Всероссийской конференции «Технологии Microsoft в теории и практике программирования» (г. Томск, 2010 г.; г. Томск, 2011 г.), 6th Central and Eastern European Software Engineering Conference (CEE-SECR) (г. Москва, 2010 г.).

По результатам исследований опубликовано 10 работ, в том числе 3 статьи в изданиях, рекомендованных ВАК, получено 2 свидетельства о государственной регистрации программ для ЭВМ.

**Личный вклад:**

1. Постановка задач исследования и разработка принципов построения инструментальной СИПД выполнены автором совместно с Н.Г. Марковым.

2. Интеграционная модель производственных данных НГДК разработана автором совместно с Н.Г. Марковым и А.В. Кудиновым.

3. Архитектура инструментальной СИПД разработана автором совместно с А.В. Кудиновым.

4. Алгоритмическое обеспечение инструментальной СИПД разработано автором. Результаты исследования эффективности предложенных алгоритмов и способа сбора и передачи технологических данных получены автором.

5. Разработка программного обеспечения инструментальной СИПД выполнена автором.

6. Разработка адаптера инструментальной СИПД, реализующего предложенный способ сбора и передачи (доставки) технологических данных АСУТП, проведена автором совместно с М.В. Копновым, А.В. Кудиновым и П.М. Острась.

7. Апробация и внедрение разработанного программного обеспечения СИПД в ОАО «Томскгазпром» выполнено совместно с А.В. Кудиновым и М.В. Копновым.

**Объем и структура работы.** Диссертация состоит из введения, четырех глав, заключения, списка использованных источников из 102 наименований и двух приложений. Объем основного текста диссертации составляет 130 страниц машинописного текста, иллюстрированного 53 рисунками и сопровождаемого 7 таблицами.

## **Содержание работы**

**В первой главе** рассматривается проблема интеграции производственных данных НГДК.

На основе результатов анализа нескольких российских НГДК описывается структура и бизнес-процессы типичной НГДК и рассматривается модель ее автоматизации. Показывается актуальность проблемы интеграции существующих в НГДК автономных ИС и комплексов программ и создания тем самым единого информационного пространства (ЕИП) компании. На основе проведенного исследования выбирается наиболее перспективный путь создания ЕИП: интеграция по данным в первую очередь автономных ИС и ком-

плексов программ, автоматизирующих производственную деятельность компании.

Рассмотрены потоки производственных данных НГДК, показаны основные трудности при их интеграции. Во-первых, это разнородность эксплуатируемых ИС и комплексов программ, во-вторых, территориальная удаленность промыслов от офисов компаний и пространственная распределенность используемых АСУТП на каждом промысле и, в-третьих, огромные объемы технологических данных как основного подмножества множества производственных данных.

Проанализированы существующие подходы и технологии интеграции производственных данных промышленных предприятий, в том числе НГДК. Сделан вывод, что существующие технологии, подходы и стандарты интеграции недостаточно эффективно решают задачу интеграции производственных данных НГДК и, соответственно, необходимо их развивать.

С учетом результатов проведенного анализа и выбранного подхода к интеграции по данным ИС и комплексов программ, эксплуатируемых в НГДК, формулируются цель исследований и задачи, решаемые в диссертационной работе.

**Вторая глава** посвящена изложению сформулированных принципов построения инструментальной системы для интеграции производственных данных (СИПД) НГДК и описанию предложенной ее архитектуры, а также созданной интеграционной модели производственных данных (ИМПД).

Даются определения участников интеграции, интеграционных процессов, схемы данных и сопоставления схем данных. *Участниками интеграции* (далее просто участниками) считаются ИС, отдельные программы, комплексы программ, базы и хранилища данных, а также АСУТП. Участники могут выступать как в роли источников, так и потребителей производственных данных в НГДК. Под *интеграционным процессом* понимается процесс передачи производственных данных от участников-источников участнику-потребителю. Под *схемой данных* понимаются любое формальное описание способа организации хранения данных с точки зрения их логической структуры, которая описывает правила хранения данных и (или) способ организации доступа к ним. Примерами схем данных являются схема базы данных и XSD (XML Schema Definition). Под *сопоставлением схем данных* будем понимать процесс определения отношения элементов двух схем (англ. mapping).

Формулируются следующие принципы построения инструментальной СИПД.

1. Принцип унифицированного федеративного доступа к данным, согласно которому каждый участник должен взаимодействовать с инструментальной системой, выступающей в роли сервера федерализации данных.
2. Принцип интеграции данных НГДК на основе интеграционной модели производственных данных.

3. Принцип адаптивности инструментальной СИПД, в соответствии с которым особенности конкретной НГДК должны легко учитываться при создании конкретной СИПД путем развития, масштабирования и адаптации инструментальной СИПД этой компании благодаря наличию в ее составе развитых инструментальных средств.

4. Принцип отдельного описания логики процессов передачи данных и описания параметров доступа к данным конкретных интегрируемых участников. Согласно ему для каждого участника необходим адаптер, предоставляющий инструментальной СИПД и, соответственно, конкретной СИПД интерфейс к его данным в виде, соответствующем ИМПД. Адаптер должен преобразовывать информацию из схемы данных участника в схему данных ИМПД.

5. Принцип сервисной ориентированности архитектуры инструментальной системы. Согласно ему транспортный уровень движения данных интегрируемых участников должен быть организован посредством веб-сервисов, а также передача данных должна быть реализована в формате XML.

6. Принцип информационной безопасности, согласно которому инструментальная система должна обладать развитыми механизмами защиты передаваемых данных, а также уметь адаптироваться к различным архитектурам безопасности интегрируемых участников и требованиям к информационной безопасности конкретных НГДК.

Рассматриваются требования, предъявляемые к разрабатываемой инструментальной системе. Среди них основными являются следующие.

1. Инструментальная СИПД должна использовать для хранения ИМПД базу данных, управляемую универсальной СУБД.

2. Для унификации взаимодействия ядра СИПД с адаптерами участников должен быть разработан формат описания запросов и ответов на запросы к объектам ИМПД.

3. Наличие в составе инструментальной СИПД инструментальных средств для упрощения процессов разработки адаптеров участников, масштабирования и адаптации такой системы к конкретной НГДК.

4. Инструментальная СИПД должна обеспечивать сбор и передачу (доставку) технологических данных АСУТП, используемых в нефтегазовой отрасли, на верхние уровни управления НГДК.

С учетом сформулированных принципов построения и требований к функциональным возможностям инструментальной СИПД была предложена архитектура такой системы (рис. 1.). Важной особенностью инструментальной системы, отличающей ее от существующих платформ интеграции, является объединение при создании ее архитектуры подхода к интеграции данных на основе интеграционной модели, базирующейся на отраслевых стандартах нефтегазовой промышленности, с принципами сервисно-ориентированной архитектуры SOA.

Основными компонентами инструментальной СИПД являются ядро и адаптеры, реализующие механизм унифицированного взаимодействия участ-

ников интеграции с ядром. Адаптер предоставляет ядру интерфейс к данным участника в виде, описанном в рамках разработанной ИМПД. Создание адаптера для подключения участника к ядру инструментальной СИПД является трудоемкой задачей, т.к. адаптер должен преобразовывать информацию из схемы данных интегрируемого участника в схему данных ИМПД. Для минимизации усилий разработчика при решении этой задачи предложено использовать алгоритмы автоматического сопоставления схем данных.

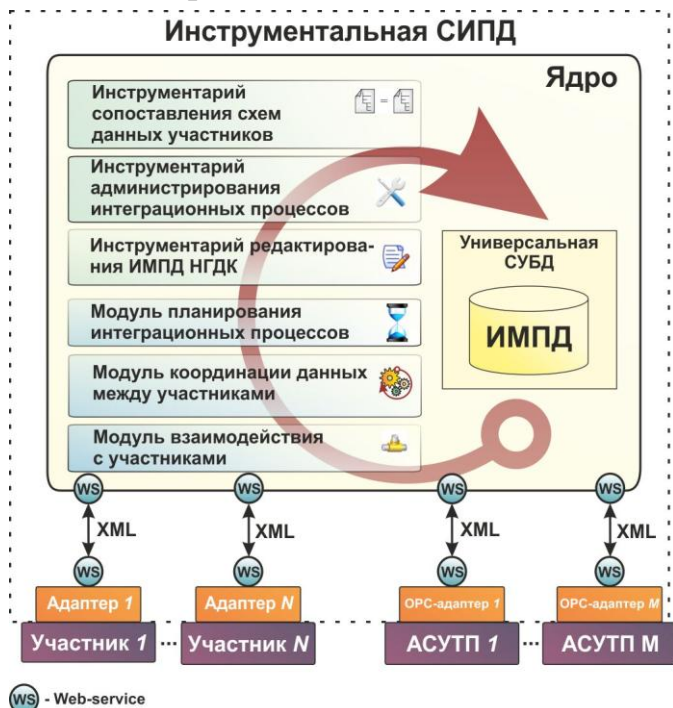


Рис. 1. Обобщенная схема архитектуры инструментальной СИПД

ности по сравнению с остальными производственными данными, для их сбора и передачи должен быть разработан отдельный ОПС-адаптер, универсальный для всех АСУТП, поддерживающих стандарт ОПС и используемых в нефтегазовой отрасли.

Основными компонентами ядра инструментальной СИПД являются (рис. 1.): набор инструментальных средств, обеспечивающих возможность построения на базе инструментальной системы конкретных СИПД для конкретных НГДК; модули, обеспечивающие выполнение основных функций инструментальной системы, и ИМПД, которая хранится в базе данных под управлением универсальной СУБД. Инструментальная система как и конкретная СИПД не хранит значения производственных показателей и их историю, а согласно первому принципу реализует функции сервера федерализации данных, позволяющие организовать интеграционные процессы внутри НГДК. Ядро инструментальной СИПД выполняет следующие основные функции:

- обеспечивает хранение ИМПД в базе данных под управлением универсальной СУБД и предоставляет средства для ее модификации;

Согласно сформулированным требованиям инструментальная система должна также обеспечивать сбор и передачу (доставку) технологических данных АСУТП. Для этого рассматривается широко используемый стандарт ОПС для сбора и передачи технологических данных, анализируются достоинства и недостатки реализующих его программных продуктов. В результате делается вывод о необходимости создания способа сбора и передачи технологических данных на основе стандарта ОПС, лишенного выявленных недостатков. Так как технологические данные имеют свои особен-

- предоставляет инструментарий для создания, редактирования, настройки, исполнения и контроля всех интеграционных процессов;
- обеспечивает обмен данными в разных режимах (по расписанию, по запросу пользователя, по изменению данных).

На основе принципов построения инструментальной СИПД и выдвинутых требований к интеграционной модели производственных данных, разработана структура такой модели (рис. 2.). ИМПД инструментальной/конкретной СИПД предлагается рассматривать как результат интеграции двух составляющих: метаданных (описания предметной области в целом) и предметной составляющей (атрибуты и отношения конкретных производственных объектов участников, описание интеграционных процессов).



Рис. 2 Структура ИМПД

Анализируются потенциальные подходы к организации хранения производственных данных в ИМПД и показываются преимущества применения в этой модели подхода «Сущность-Атрибут-Значение» (англ. Entity-Attribute-Value model, EAV) перед другими. Проектируются структуры для хранения данных ИМПД на концептуальном уровне.

Показывается, что наиболее зрелым стандартом нефтегазовой отрасли, имеющим в своем составе описание производственных объектов, является PRODML, т. к. схема данных, предлагаемая этим стандартом, покрывает практически весь производственный процесс с момента добычи углеводородного

сырья (УВС), подготовки товарной продукции до момента ее реализации. Классы и их атрибуты, описывающие производственные объекты в составе стандарта PRODML, взяты за основу метасоставляющей ИМПД. Структура интеграционной модели предусматривает возможность дополнения основного набора классов и их атрибутов, это и определяет гибкость модели в соответствии с требованиями конкретной НГДК и ее СИПД.

Проведен анализ эффективности предложенной интеграционной модели данных и ряда других моделей данных, используемых для описания производственных данных в современных НГДК. По его результатам сделан вывод, что положенный в основу ИМПД подход позволяет проводить расширение набора классов объектов, числа их атрибутов, отношений между классами и между объектами без структурных изменений схемы данных.

**В третьей главе** описывается алгоритмическое обеспечение инструментальной СИПД.

Разработан алгоритм координации данных между участниками интеграции по данным, в соответствии с ним исполняются интеграционные процессы инструментальной системой.

Описан разработанный алгоритм PSM, обеспечивающий автоматическое сопоставление схем данных и позволяющий минимизировать усилия разработчика при подключении участников к ядру инструментальной СИПД. Этот алгоритм включает в себя ряд алгоритмов сопоставления схем данных: синтаксический, семантический, структурный алгоритмы и алгоритм сопоставления схем данных на основе типов данных. Алгоритм PSM осуществляет поиск соответствий между двумя схемами данных, представленными в формате XSD, и относится к классу алгоритмов, использующих при сопоставлении только схему данных, т. е. имена элементов, типы данных и структурные свойства элементов схем. Укрупненное описание алгоритма PSM следующее.

**Начало.**

**Шаг 1.** Разбить каждую строку названия элементов схем данных на подстроки по набору правил.

Примечание: необходимость разбиения названий элементов обусловлена тем, что они могут быть составными (Camel-стиль, Pascal-стиль и т. п.).

**Шаг 2.** Рассчитать соответствия между списками подстрок названий элементов и их синонимов из словаря с использованием синтаксических алгоритмов (Левенштейна, N-грамм и комбинированного).

**Шаг 3.** Вычислить соответствия между элементами схем по типам данных.

**Шаг 4.** Вычислить соответствия между элементами схем данных на основе структурных зависимостей. На этом шаге используются результаты, полученные на шаге 2 и шаге 3. Детализированный шаг 4 в виде схемы алгоритма приведен на рис. 3.

**Шаг 5.** Выбрать соответствия с максимальными весами между элементами исходных схем данных на основе матриц результатов, полученных на шаге 2, шаге 3 и шаге 4.

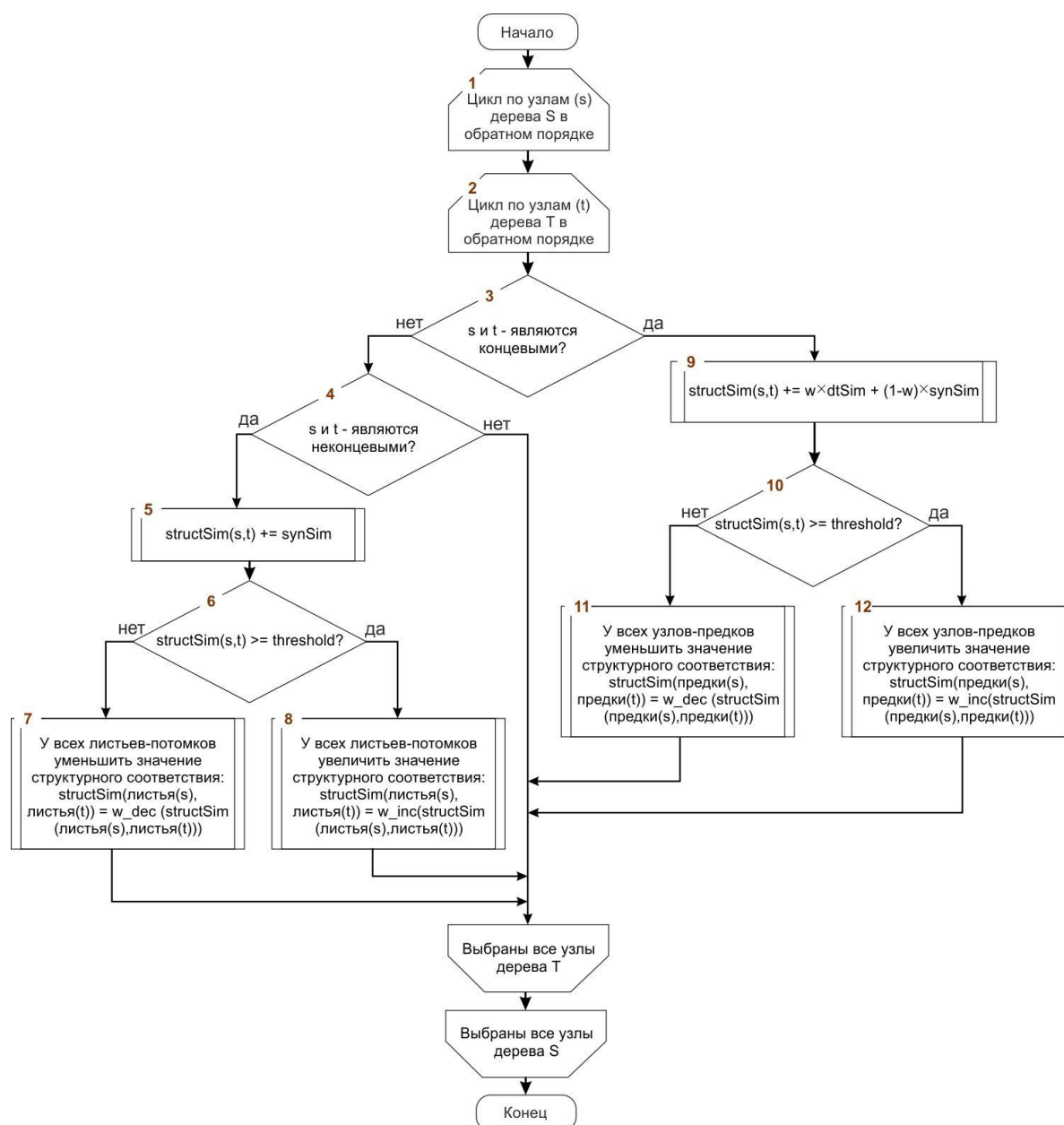
**Конец.**

В литературе не представлены алгоритмы структурного сопоставления схем данных, поэтому был разработан оригинальный эвристический алгоритм (рис. 3.). Учтено, что схему данных можно представить графом, но чаще всего это ациклический граф (не содержит циклов), т. е. дерево. Введены понятия для пояснения предлагаемого алгоритма: узлы с нулевой степенью исхода (из которых не исходит ни одна дуга), называемые концевыми вершинами или листьями; узел ветвления (который имеет входящие и исходящие дуги) – неконцевой узел; дерево с отмеченным узлом, называемое корневым деревом.

Предложенный алгоритм структурного сопоставления деревьев схем данных базируется на следующих предположениях:

- листья двух деревьев соответствуют друг другу в большей степени, если их предки соответствуют друг другу синтаксически и по типу данных;
- неконцевые узлы деревьев соответствуют друг другу структурно, если они соответствуют синтаксически и их корневые деревья соответствуют друг другу;

- неконцевые узлы деревьев соответствуют друг другу структурно, если их листья имеют высокую степень соответствия, даже если их непосредственные дети слабо соответствуют друг другу; это связано с тем, что листья обычно имеют более высокий семантический смысл по сравнению с неконцевыми узлами.



Изменение значения структурного соответствия  $w_{inc}(\lambda)$  и  $w_{dec}(\lambda)$  (шаги 7, 8, 11, 12 на рис. 3) рассчитывается на основе функции вида:

$F(X) = X / (1 + |X|)$ , имеющей гарантированное предельное значение, равное единице. Значения после приращения и уменьшения структурного веса узла вычисляются по следующим формулам:

$$w\_inc(\lambda) = cur + c \times \lambda \times (1 - cur) / (1 + c \times \lambda \times (1 - cur))$$

$$w\_dec(\lambda) = cur - c \times \lambda \times (1 - cur) / (1 - c \times \lambda \times (1 - cur)),$$

где  $\lambda$  – вес соответствия, на основании которого принято решение об увеличении или уменьшении значения структурного веса текущего узла,  $c$  – коэффициент скорости приращения веса,  $cur$  – значение веса текущего узла.

Исследование эффективности алгоритма PSM выполнено на ряде тестовых схем данных, а также на наборе реальных схем данных, описывающих результаты гидродинамических исследований скважин и используемых популярными в нефтегазовой отрасли ИС и комплексом программ (КГСУ «Магистраль-Восток», ИС «BASPRO» и программный продукт фирмы Schlumberger «InterACT»). Оценка алгоритма PSM с точки зрения качества сопоставления схем данных (F-мера) получена в сравнении с оценкой комбинации алгоритмов широко используемой программной среды COMA+ (EditDistance, Synonym, NameType и NamePath), позволяющей комбинировать различные алгоритмы сопоставления схем данных. Для оценки качества использована F-мера:

$$F\text{-мера} = \text{точность} \times \text{полнота} / (0.5 \times \text{точность} + 0.5 \times \text{полнота}),$$

где *точность* – доля найденных истинных положительных соответствий среди всех соответствий, *полнота* – доля найденных истинных положительных соответствий среди всех истинных соответствий.

На основе результатов исследования выявлено, что средний выигрыш алгоритма PSM по качеству сопоставления схем по сравнению с алгоритмами среды COMA+ на выбранных наборах составил 10,1% (рис. 4.).

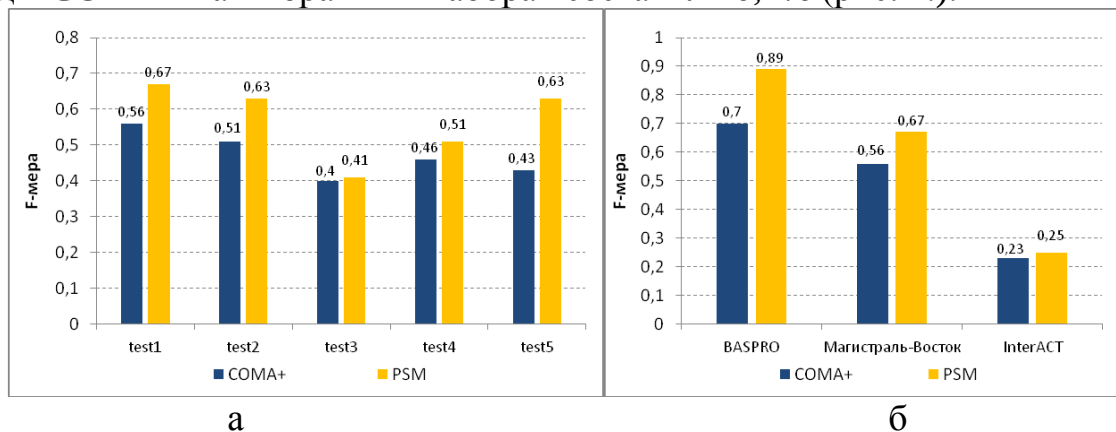


Рис. 4. Результаты оценки качества сопоставления (F-мера) алгоритмов на исследуемых схемах данных: а – тестовые схемы данных; б –схемы данных реальных систем

Разработан алгоритм автоматического сопоставления объектов ИМПД, позволяющий качественно улучшить настройку интеграционных процессов инструментальной/конкретной СИПД.

Предложен оригинальный способ сбора и передачи (доставки) технологических данных на верхние уровни управления НГДК от используемых в

компании АСУТП. Ключевыми особенностями его, обеспечивающими информационную защиту передаваемых данных, является использование сервисно-ориентированной и двухзвенной архитектуры программного средства доставки данных, реализующего способ. Использование сервисно-ориентированной архитектуры, элементом которой являются веб-сервисы, позволяет выполнить все предъявляемые в НГДК требования по информационной безопасности к средствам доставки технологических данных. Протоколом обмена данными при использовании веб-сервисов является SOAP (Simple Object Access Protocol), использующий формат XML, поэтому имеется возможность фильтрации всего трафика от АСУТП в СИПД межсетевым экраном. Предложенная двухзвенная архитектура программного средства сбора и передачи технологических данных позволяет передающей части этого средства находиться в технологической локальной вычислительной сети (ЛВС) вместе с OPC-сервером, принимающим данные от АСУТП, а другой части в виде веб-сервиса, выполняющего функцию приемника данных в XML-формате, находиться вместе с OPC-адаптером, отвечающим за доставку данных потребителям, в офисной ЛВС. Принимающий данные веб-сервис представляет собой «пассивный» интерфейс, т. е. все действия инициируются передающей частью, расположенной в технологической ЛВС. Это удовлетворяет самым жестким требованиям информационной безопасности в НГДК, т. к. исключает интерфейс доступа к данным АСУТП в технологическую сеть.

Средством обеспечения надежности предлагаемого способа является механизм буферизации, который позволяет исключить потери данных при отсутствии связи между технологической и офисной вычислительными сетями (по аналогии со спецификацией OPC HDA). Для реализации этого механизма может быть использована БД под управлением СУБД или организован файловый буфер, накапливающий технологические данные до восстановления соединения между сетями.

Реализация предлагаемого способа предполагает использование ряда подходов, позволяющих уменьшить объем пакетов передаваемых технологических данных.

Для проведения исследования эффективности предложенного способа он реализован в виде программы, относящейся к классу OPC-клиентов. Проведено исследование его производительности при доставке данных. Аналогичные исследования выполнены также для OPC Easy Connect Suite компании Softing и XML-Data Access Gateway компании Advosol, являющихся широко распространенными OPC-клиентами, реализующими спецификацию OPC XML DA.

Анализ результатов исследования (рис. 5.) показал, что разработанный OPC-клиент дает значительный выигрыш в производительности по сравнению с аналогами. Например, при пропускной способности канала передачи данных 1 Мбит/с разработанный OPC-клиент может доставить 8 тысяч значений (каждое значение – кортеж атрибутов) технологических параметров за

1 минуту против 4–6 тысяч значений, обеспечиваемых OPC-клиентами Easy Connect Suite и Data Access Gateway.

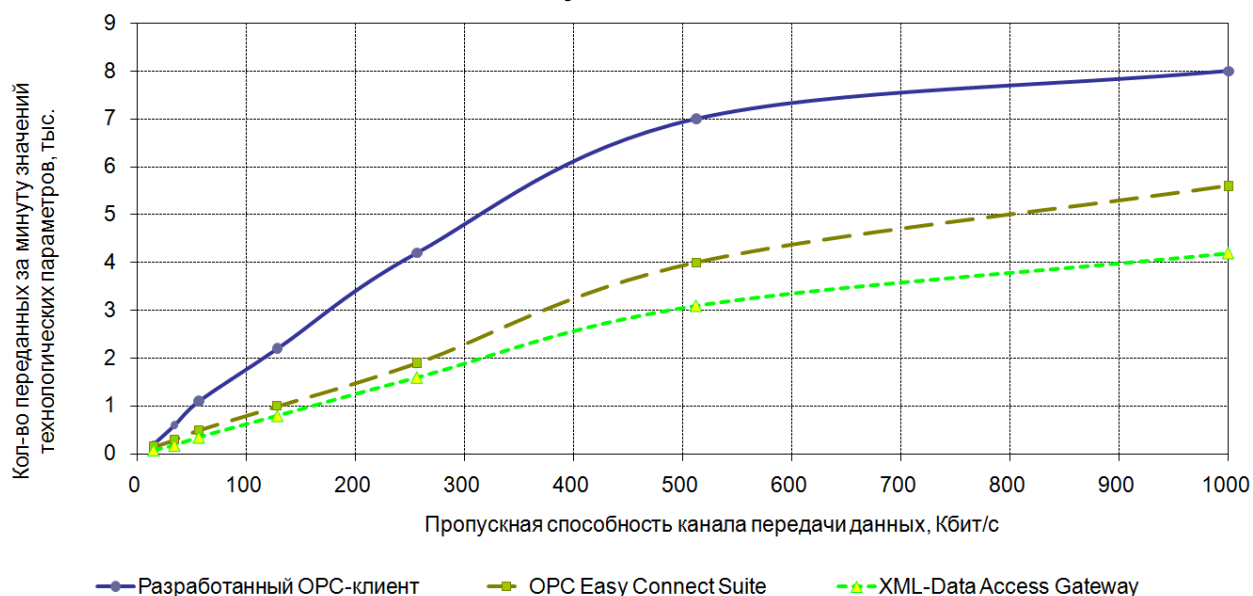


Рис. 5. Зависимость количества переданных за минуту значений технологических параметров от пропускной способности канала передачи данных

Этот выигрыш объясняется следующими факторами. Во-первых, в OPC-клиенте реализована не XML DA спецификация, а оригинальный интерфейс доступа к технологическим данным, XML-схемы которого более простые, и, следовательно, размер передаваемых пакетов тех же данных значительно снижен. Во-вторых, значения технологических параметров в передаваемых пакетах группируются по времени, т.к. при передаче значений параметров в асинхронном режиме (этот режим наиболее распространен в нефтегазовой отрасли), атрибут «время поступления значения» большинства значений одинаков. В-третьих, уменьшен объем пакетов за счет использования коротких имен тегов в пакетах. В-четвертых, в пакетах передаются не сами значения параметров, а разница текущего и предыдущего значений параметра.

**В четвертой главе** описывается программная реализация разработанной инструментальной СИПД и результаты ее апробации и внедрения в НГДК ОАО «Томскгазпром» созданной на ее основе СИПД.

В качестве программных платформ для разработки инструментальной СИПД выбраны Microsoft .NET Framework, WCF (Windows Communication Framework) и LINQ (Language Integrated Query). В качестве средства разработки на этих платформах использована среда Microsoft Visual Studio 2008 и язык программирования C#. Учитывая возможность тесной интеграции с выбранными средствами разработки, в качестве СУБД выбрана Microsoft SQL Server 2008. OPC-клиент разработан на языке Delphi в среде CodeGear Delphi 2007.

На основе концептуальной схемы данных, представленной во второй главе, разработана физическая схема данных ИМПД.

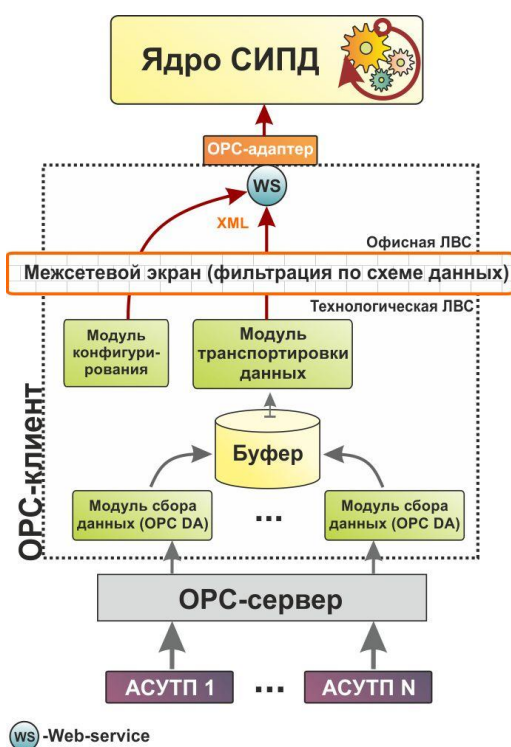


Рис. 6. Архитектура ПО разработанного OPC-клиента

При внедрении СИПД осуществлена интеграция по данным под-лежала КГСУ «Магистраль–Восток» с двумя специализированными ИС: «BASPRO» (позволяет обрабатывать и интерпретировать геолого-геофизические данные, строить геологические модели) и «OISPipe» (модели-рование промысловых нефте- и газосборных сетей) (рис.7.). КГСУ «Магист-раль-Восток» – это система класса MES, внедрена в компании для автоматизации деятельности большинства производственных служб. Основные произ-водственные объекты и их атрибуты, по которым организована передача производственных данных между этими системами следующие:

- скважины – давление (буферное, затрубное), температура (линейная, устьевая), состояние скважины, способ эксплуатации скважины, время работы/простоя, обводненность, компонентный состав сырья, результаты гидродинамических исследований;
- пласты – плотность воды, плотность нефти, газосодержание;
- пластопересечения (эксплуатационные пласты) – добыча/выпуск УВС, добыча воды;
- и т.д.

КГСУ «Магистраль-Восток» интегрировалась также по технологиче-ским данным с АСУТП кустов скважин и установок подготовки газа и газо-вого конденсата (Мыльджинское и Северо-Васюганское газоконденсатные месторождения), а также установки по подготовки нефти (Казанское место-рождение).

Это обеспечило сбор и передачу (доставку) данных, представляющих собой значения параметров технологических процессов (дебит, давление,

Описана детальная структура ПО ин-струментальной СИПД. Подробно рас-смотрены основные программные модули, входящие в ее состав: редактирования ИМПД, сопоставления схем данных, на-стройки интеграционных процессов, пла-нирования интеграционных процессов, взаимодействия ядра СИПД с участника-ми, координации данных между участни-ками и формирования запросов к адапте-рам участников.

Разработана архитектура ПО OPC-клиента, созданного на основе предложен-ного способа сбора и передачи (доставки) технологических данных и подключаемого к ядру СИПД с помощью универсального OPC-адаптера (рис. 6.).

Приведены результаты апробации ин-струментальной системы и внедрение соз-данной на ее основе СИПД ОАО «Томск-

температура и т. д.) через заданные промежутки времени, в КГСУ «Магистраль-Восток».

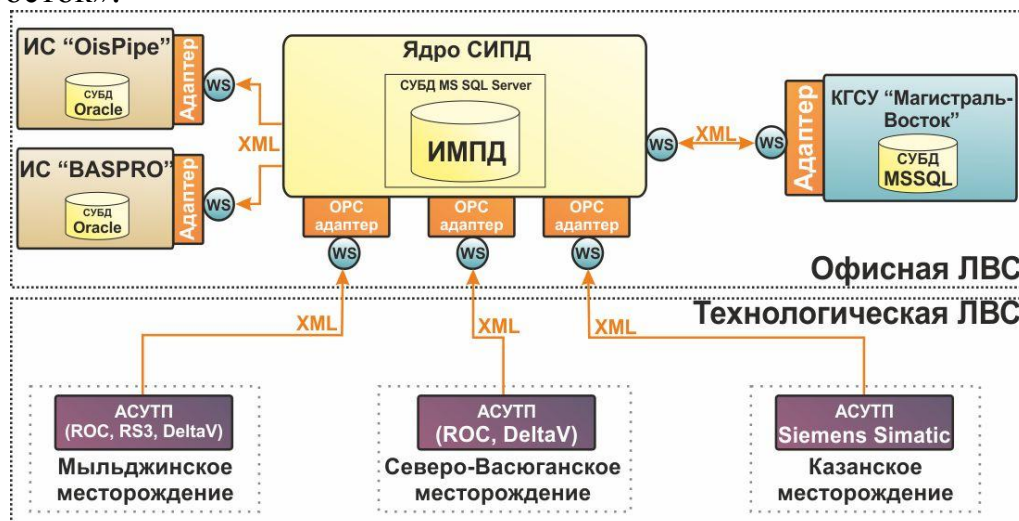


Рис. 7. Фрагмент СИПД ОАО «Томскгазпром»

При помощи OPC-клиентов (на рис. 7 изображены только OPC-адаптеры) доставляются данные из 8 АСУТП различного назначения с трех месторождений ОАО «Томскгазпром» и с Лугинецкого узла учета и сдачи стабильного конденсата и нефти в магистральный нефтепровод «Александровское – Томск – Анжеро-Судженск», всего более 6600 технологических параметров. Апробация СИПД при сборе и передаче технологических данных компании показала, что обеспечивается доставка примерно 880 тысяч значений параметров в сутки. Использование механизма буферизации, реализованного в OPC-клиенте, в течение последних трех лет позволило не потерять за время перерывов связи 1,3 млн. значений параметров. Кроме того, внедрение разработанного OPC-клиента позволило ОАО «Томскгазпром» в 2008 г. успешно пройти аудит комиссии по информационной безопасности ОАО «Газпром». Как показала практика, адаптация инструментальной СИПД (масштабирование СИПД ОАО «Томскгазпром») с целью организации сбора и доставки значений технологических параметров с вводимых в эксплуатацию новых АСУТП, производится в короткие сроки.

Результаты внедрений подтверждают эффективность разработанных алгоритмических и программных средств инструментальной/конкретной СИПД и их практическую значимость.

**В заключении** приведены основные полученные в диссертационной работе результаты и сделанные выводы.

**В приложения** вынесены акты о внедрении полученных результатов и свидетельства о регистрации программ для ЭВМ.

## Основные результаты и выводы

В ходе выполнения диссертационной работы были получены следующие основные научные и практические результаты и сделаны выводы.

1. Проведен анализ проблемы интеграции эксплуатируемых в НГДК информационных систем и комплексов программ производственного назна-

чения. Показано, что наиболее перспективным для решения практических задач в рамках решения этой проблемы является подход, когда ИС и комплексы программ НГДК интегрируются по производственным данным. По результатам анализа поставлена цель создания алгоритмического и программного обеспечения инструментальной системы для интеграции производственных данных ИС и комплексов программ, производственного назначения, эксплуатируемых в нефтегазодобывающих компаниях.

2. Предложены принципы создания и разработана архитектура инструментальной СИПД.

3. На основе стандарта PRODML разработана интеграционная модель производственных данных, позволяющая гибко описывать производственные объекты добывающих компаний нефтегазовой отрасли.

4. Разработано алгоритмическое обеспечение инструментальной СИПД. Предложен алгоритм PSM, решающий задачу автоматического сопоставления схем данных и позволяющий упростить подключение участников интеграции к ядру инструментальной СИПД. Проведено исследование эффективности алгоритма PSM и комбинации алгоритмов из широко распространенной среды СОМА+. Результаты исследования показали, что алгоритм PSM обеспечивает более высокое качество сопоставления схем данных.

5. Разработан оригинальный способ сбора и передачи (доставки) технологических данных на верхние уровни управления компанией, базирующийся на стандарте OPC и принципах сервисно-ориентированной архитектуры. Способ обеспечивает высокий уровень информационной безопасности и надежности. Проведен анализ производительности его программной реализации в виде OPC-клиента. Сделан вывод о том, что разработанный OPC-клиент обеспечивает более высокую производительность по сравнению с OPC-клиентами, широко используемыми сегодня в компаниях нефтегазовой отрасли.

6. Создано программное обеспечение инструментальной СИПД. Объем исходного кода разработанных программных средств составляет более 17 000 строк на языках C#, Delphi, T-SQL и PL/SQL. Часть программ имеет государственную регистрацию.

7. Проведены апробация разработанного алгоритмического и программного обеспечения инструментальной СИПД и внедрение созданной на ее основе конкретной СИПД в НГДК ОАО «Томскгазпром». Результаты апробации и внедрения подтверждают эффективность разработанных архитектуры, алгоритмических и программных средств инструментальной системы и СИПД для ОАО «Томскгазпром».

### **Основные публикации по теме диссертации**

#### **Статьи, опубликованные в изданиях из перечня ВАК:**

1. Вейбер В.В., Богдан С.А., Кудинов А.В., Марков Н.Г. Концепция построения платформы для интеграции производственных данных нефтегазо-

добывающей компании // Известия Томского политехнического университета. — 2011. — Т.318. — № 5. — С. 126-131.

2. Вейбер В.В., Кудинов А.В., Марков Н.Г. Алгоритм сопоставления схем данных информационных систем нефтегазодобывающего предприятия // Известия Томского политехнического университета. — 2011. — Т.319. — № 5. — С. 178-183.

3. Вейбер В.В., Кудинов А.В., Марков Н.Г. Задача сбора и передачи технологической информации распределенного промышленного предприятия // Известия Томского политехнического университета. — 2011. — Т.319. — №5. — С. 151-156.

**Статьи, тезисы докладов на междунар. и всеросс. конференциях:**

4. Вейбер В.В., Ковин Р.В. Разработка OPC-клиента на основе веб-сервисов // Молодежь и современные информационные технологии: Сб. трудов VI Всеросс. научно-практической конф. студентов, аспирантов и молодых ученых. — Томск: Изд-во СПБ Графикс. — 2008. — С. 461–462.

5. Вейбер В.В., Марков Н.Г. Сервисно-ориентированная интеграционная платформа на базе WCF/WFF для построения единого информационного пространства нефтегазодобывающего предприятия // Сб. трудов VI Всеросс. научно-практической конф. «Технологии Microsoft в теории и практике программирования». — Томск: Изд-во ТПУ. — 2009. — С. 86–88.

6. Вейбер В.В., Марков Н.Г. Использование сервисно-ориентированной архитектуры при построении систем управления нефтегазодобывающим предприятием // Молодежь и современные информационные технологии: Сб. трудов VII Всеросс. научно-практической конф. студентов, аспирантов и молодых ученых. — Томск: Изд-во СПБ Графикс. — 2009. — С. 20–21.

7. Вейбер В.В., Кудинов А.В. Программное решение для интеграции информационных систем промышленного предприятия // Современные техника и технологии: Труды XV Междун. научно-практической конф. студентов, аспирантов и молодых ученых — Томск: Изд-во ТПУ. — 2009. — Т. 2. — С. 181–183.

8. Вейбер В.В., Кудинов А.В., Марков Н.Г. Использование метамодели предметной области для интеграции производственных данных нефтегазового предприятия // Сб. трудов VII Всеросс. научно-практической конф. «Технологии Microsoft в теории и практике программирования». — Томск: Изд-во ТПУ. — 2010. — С. 67–70.

9. Veyber V., Kudinov A., Markov N. Model driven approach for oil & gas information systems and applications integration // Proceedings of 6th Central and Eastern European Software Engineering Conference (CEE-SECR 2010), — Moscow. — ISBN 978-1-4577-0605-9. — 2010. — P. 156–162.

10. Вейбер В.В., Кудинов А.В. Платформа интеграции производственных данных нефтегазодобывающей компании // Сб. трудов VIII Всеросс. научно-практической конф. «Технологии Microsoft в теории и практике программирования». — Томск: Изд-во ТПУ. — 2011. — С. 98–100.

Подписано к печати 17.11.2011 г. Формат \_\_\_\_\_  
\_\_\_\_\_ Усл.печл. \_\_\_\_\_ Уч.-издл. \_\_\_\_\_  
Заказ \_\_\_\_\_ Тираж 100 экз.

Издательство ТПУ. 634050, г. Томск, пр. Ленина, 30  
Тел./факс: 8(3822)56-35-35, [www.tpu.ru](http://www.tpu.ru)