
РАЗРАБОТКА МОДЕЛИ ХРАНИЛИЩА ДАННЫХ ИННОВАЦИОННОГО ПРЕДПРИЯТИЯ ПРИ ВЫСШЕМ УЧЕБНОМ ЗАВЕДЕНИИ

**Ф.О. Федин, С.В. Чискидов,
Е.Н. Павличева**

Кафедра прикладной информатики
Московский городской педагогический университет
2-й Тульский переулок, 4, Москва, Россия, 115191

В работе выполнено моделирование бизнес-процессов малого инновационного предприятия ИТ-профиля, функционирующего при высшем учебном заведении, а также проектирование логического и физического уровней модели хранилища данных этого предприятия. Работа выполнена с использованием специализированных CASE-средств фирмы Computer Associates — ERWin Process Modeler и ERWin Data Modeler.

Ключевые слова: малое инновационное предприятие, хранилище данных, Dimensional, SQL-запрос, таблица фактов, таблица измерений.

Повышение эффективности всех уровней и форм образования за счет внедрения информационных и телекоммуникационных технологий неоднократно обосновано в результате многочисленных педагогических исследований. В соответствии со вступившим в силу Законом от 2 августа 2009 г. № 217-ФЗ высшие образовательные учреждения могут становиться учредителями малых инновационных предприятий (МИП) и таким образом осуществлять практическое применение результатов интеллектуальной деятельности, права на которые принадлежат данным учебным заведениям. Это стимулирует научную деятельность, открывает простор молодежной инициативе, способствует созданию конкурентоспособной инновационной продукции для различных отраслей экономики, включая сферу информационных систем и технологий [1].

Выполненный авторами анализ деятельности МИП ИТ-профиля, функционирующего при высшем учебном заведении и занимающегося проектированием, разработкой и тиражированием программных продуктов (игровых приложений) для мобильных устройств, позволил разработать функциональную модель его деятельности. Для построения модели был использован графический язык описания бизнес-процессов, составляющий основу методологии IDEF0 [2] и реализованный в среде моделирования и анализа бизнес-процессов СА ERWin Process Modeler [3].

Процесс создания модели деятельности МИП ИТ-профиля, функционирующего при вузе, был начат с определения контекста (построения контекстной диаграммы верхнего уровня), т.е. наиболее абстрактного уровня описания деятельности МИП в целом (рис. 1).

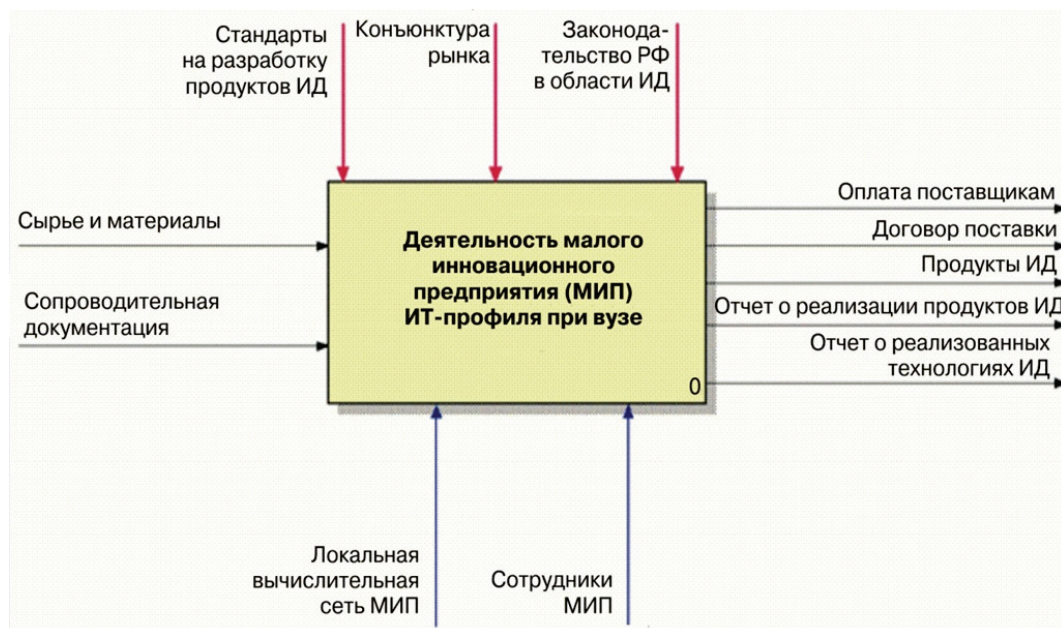


Рис. 1. Контекстная диаграмма верхнего уровня

Входными данными (стрелки слева) являются закупаемое сырье и материалы, используемые для записи, оформления и упаковки разрабатываемых программных продуктов, а также сопроводительная документация к ним. В качестве управляющих факторов (стрелки сверху) выступают существующие стандарты на разработку продуктов инновационной деятельности (ИД), результаты анализа конъюнктуры рынка (уровня спроса и предложения, цен, покупательских предпочтений, деятельности конкурентов и партнеров и др.) и законодательство РФ в области ИД. Исполнительными механизмами (стрелки снизу) являются локальная вычислительная сеть МИП и сотрудники этого предприятия (преподаватели, студенты, технический персонал и др.). Выходные данные деятельности МИП (стрелки справа) — это оплата поставщикам за поставляемое сырье и материалы, заключенные договора поставки, произведенные продукты ИД, а также различные отчеты, например отчеты о реализованных технологиях ИД.

На рис. 2 представлена диаграмма декомпозиции первого уровня (контекстная диаграмма уровня А0) модели деятельности МИП ИТ-профиля, функционирующего при высшем учебном заведении. Эта диаграмма включает четыре вида работ: выполнение закупки сырья и материалов; организация складского учета закупленного сырья и материалов, а также произведенных и ожидающих реализации продуктов ИД; непосредственно производство продуктов ИД; выполнение реализации продуктов ИД.

Детализация работ по выполнению закупки сырья и материалов, организации складского учета, производства продукта ИД и реализации продукта ИД показана на диаграммах декомпозиции, представленных на рис. 3—6.

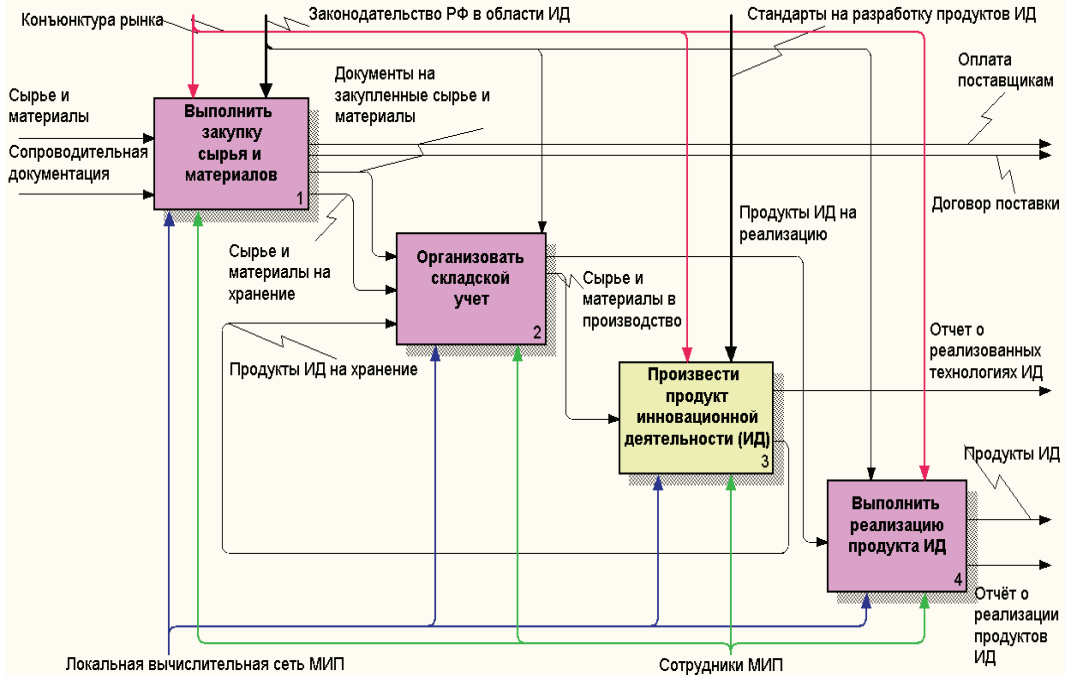


Рис. 2. Контекстная диаграмма уровня А0

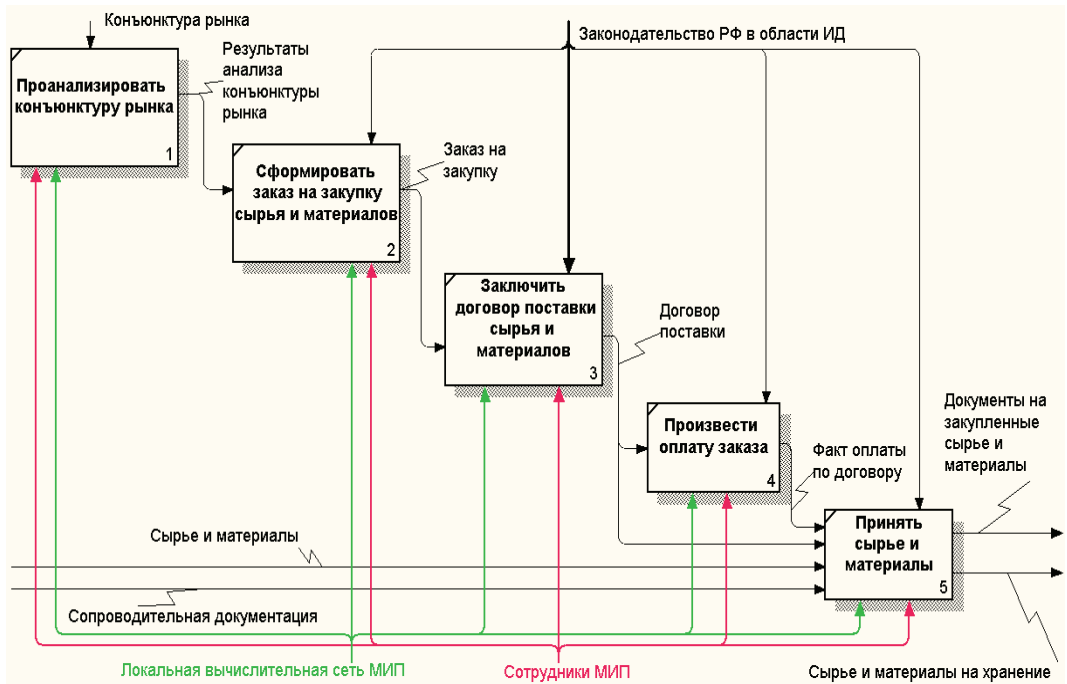


Рис. 3. Диаграмма декомпозиции работы «Выполнить закупку сырья и материалов»

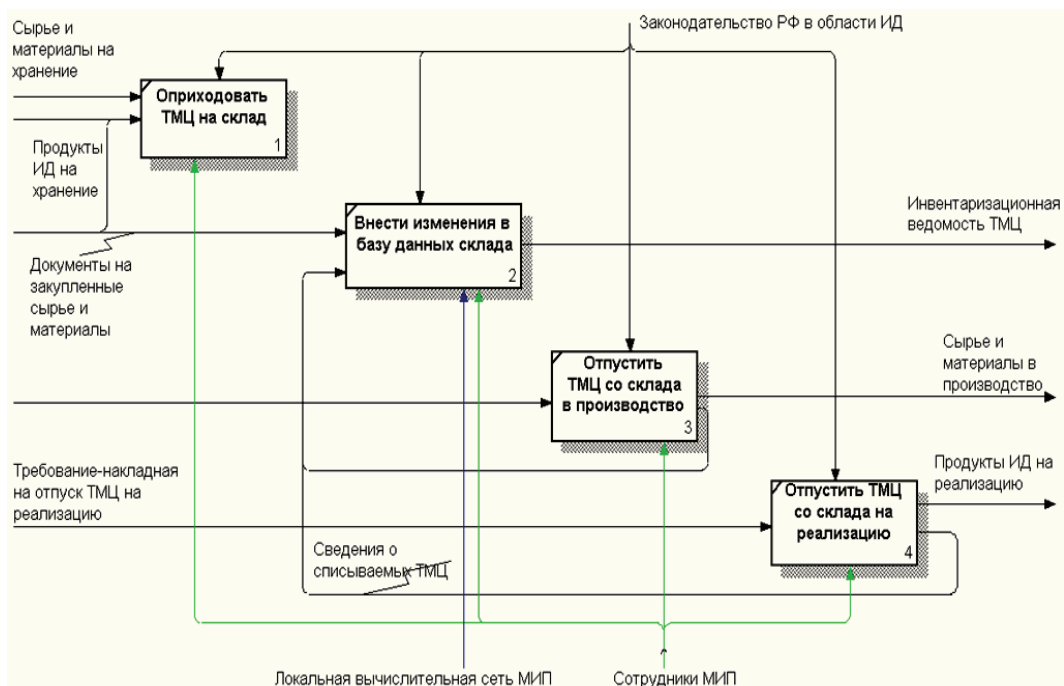


Рис. 4. Диаграмма декомпозиции работы «Организовать складской учет»

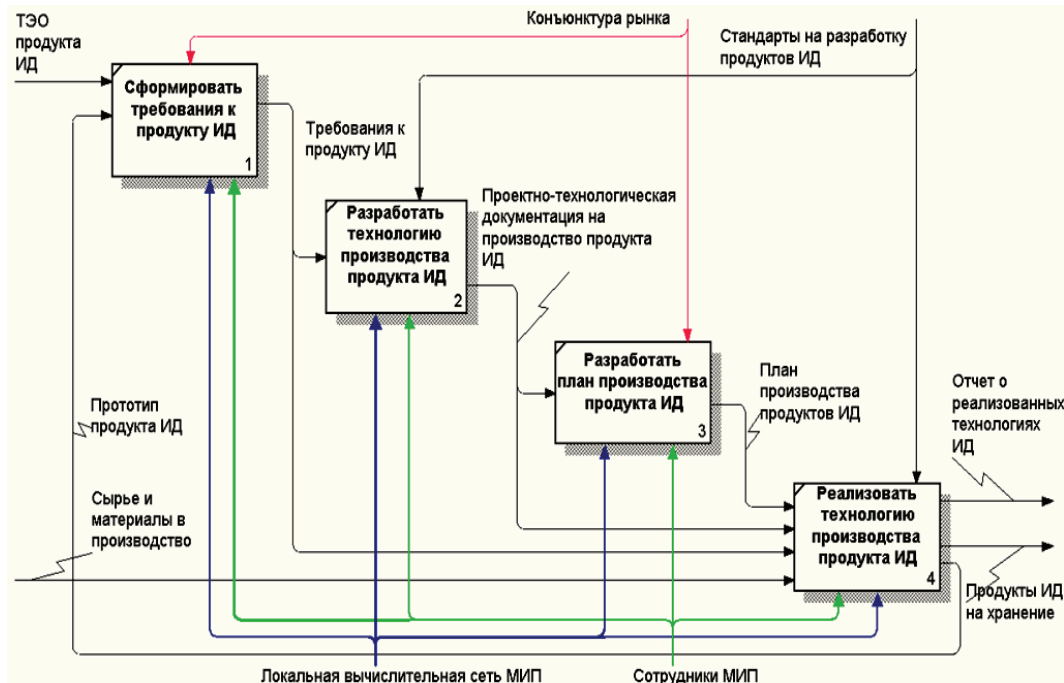


Рис. 5. Диаграмма декомпозиции работы «Произвести продукт инновационной деятельности»

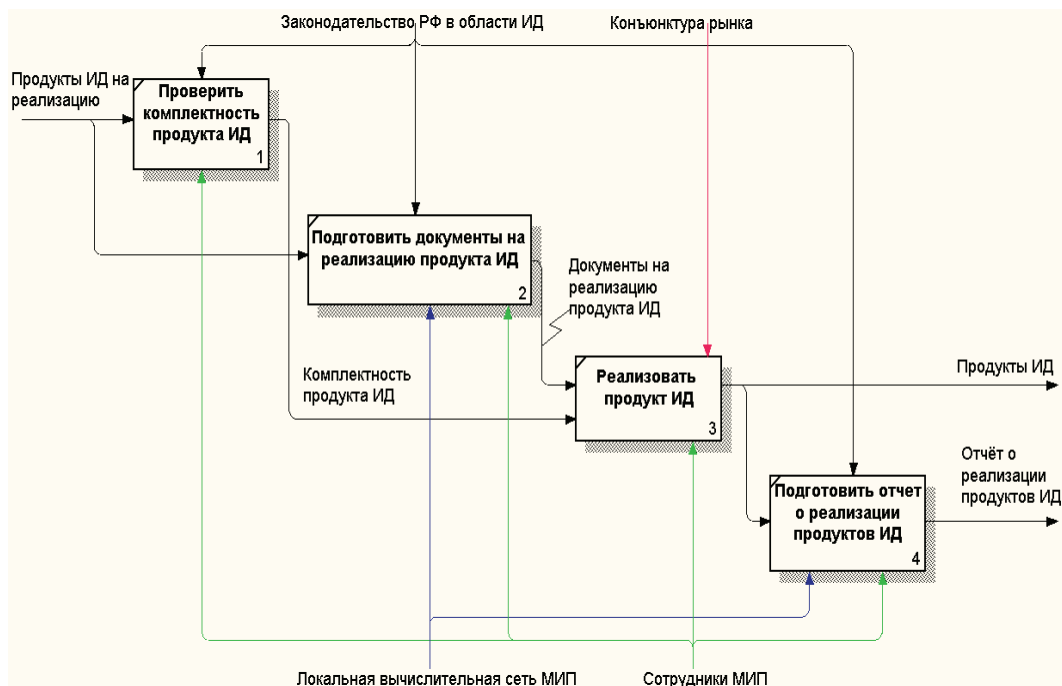


Рис. 6. Диаграмма декомпозиции работы «Выполнить реализацию продукта инновационной деятельности»

Одним из сдерживающих элементов на пути внедрения и устойчивого развития МИП ИТ-профиля в вузах является слабая информационно-аналитическая поддержка сотрудников таких предприятий, обусловленная следующими основными факторами [4; 5]: различные типы, форматы и кодировки данных о деятельности МИП; недостаточная структурированность данных, существенно затрудняющая выполнение их консолидации и сопоставления; высокая вероятность появления в данных ошибок, аномальных значений, противоречий, дубликатов, пропусков; отсутствие автоматической поддержки целостности данных; невозможность прослеживания истории изменений в данных; отсутствие возможности для конечного пользователя получать нужные данные за нужный хронологический период в нужных разрезах и форматах для формирования отчетов.

Анализ возможных механизмов, позволяющих максимально снизить или совсем исключить влияние указанных выше факторов, позволил сделать вывод о том, что наиболее подходящим из них является внедрение в деятельность МИП ИТ-профиля специализированного хранилища данных. К такому хранилищу предъявляются следующие основные требования: накопление необходимой хронологической информации о деятельности МИП ИТ-профиля и управление ее хранением; хранение как детализированных, так и обобщенных (агрегированных) данных деятельности МИП; единый формат хранения и обработки данных; доступ к накопленным за достаточно длительный период времени данным с соблюдением их хронологии; поддержка нерегламентированных сложных запросов,

формируемых «на лету» и построение OLAP-отчетов. В целях реализации указанных требований, с помощью инструментального средства ERWin Data Modeler, авторами была разработана модель хранилища данных МИП ИТ-профиля, имеющая три формы представления данных на логическом уровне: диаграмму сущность-связь (ER-диаграмму), модель основанную на ключах (КВ-модель) и полную атрибутивную модель (ФА-модель).

ER-диаграмма (рис. 7) включает сущности и связи между экземплярами этих сущностей.

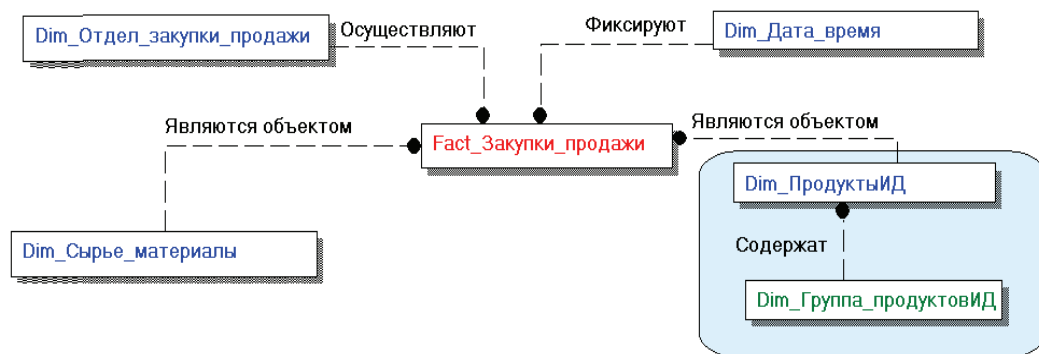


Рис. 7. ER-диаграмма хранилища данных МИП ИТ-профиля

КВ-модель (рис. 8) более подробно представляет данные и содержит сущности, связи между экземплярами сущностей, а также ключевые атрибуты в составе первичных ключей.

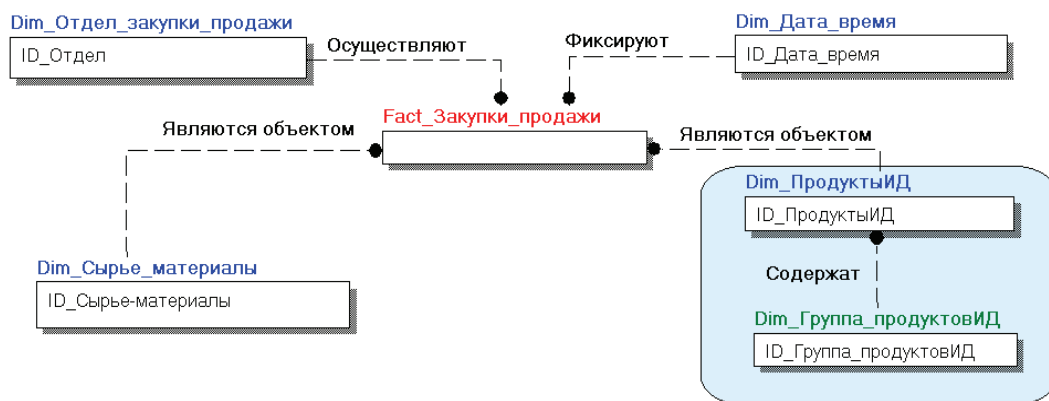


Рис. 8. КВ-модель хранилища данных МИП

Наиболее детальное представление структуры данных содержит ФА-модель (рис. 9), в которую вошли все родительские и дочерние сущности, связи между экземплярами сущностей, а также ключевые и неключевые атрибуты.

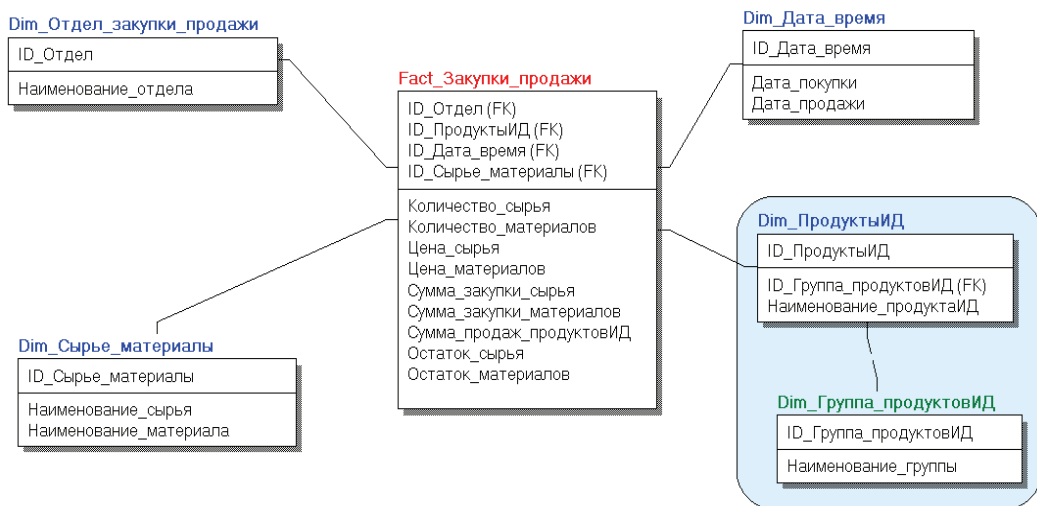


Рис. 9. FA-модель хранилища данных МИП

На физическом уровне представления данных были разработаны трансформационная (Т-модель) и размерная (Dimensional) модели хранилища данных. Т-модель (рис. 10) содержит таблицы и имена столбцов, для которых определены: типы данных в формате выбранной СУБД (Firebird); ограничения на уникальность значений столбцов, являющихся первичными ключами; ограничения на неопределенные значения столбцов; ограничения на ввод данных в столбец; ограничения внешнего ключа.

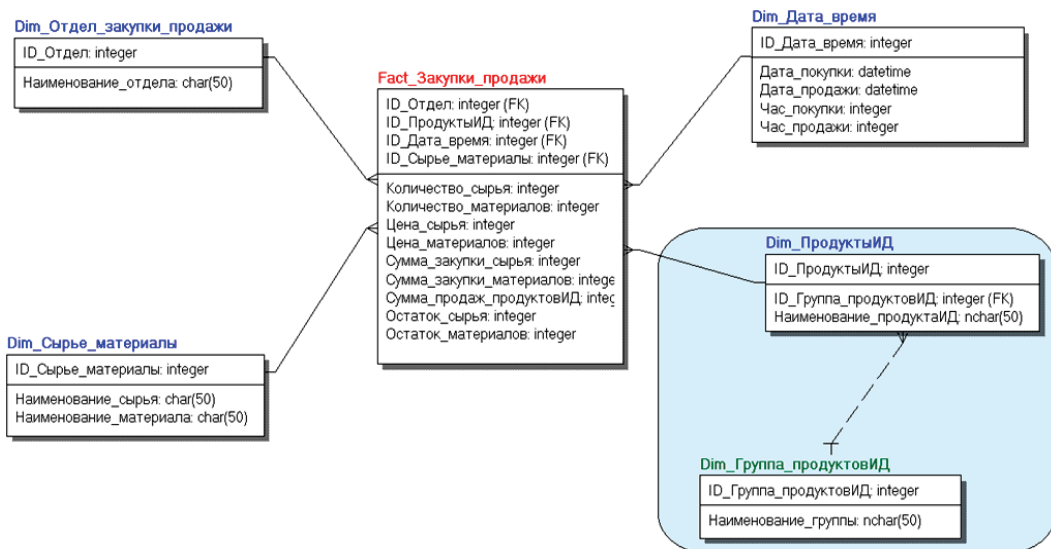


Рис. 10. Т-модель хранилища данных МИП

Моделирование Dimensional сходно с моделированием связей и сущностей для реляционной модели, но отличается целями. Реляционная модель акценти-

руется на целостности и эффективности ввода данных, а размерная модель (Dimensional) — на эффективности просмотра данных и выполнении сложных SQL-запросов [3].

При разработке размерной модели авторами был использован стандарт, называемый схемой «снежинка» (рис. 11), обеспечивающий высокую скорость выполнения запроса посредством денормализации и разделения данных.

Создать универсальную денормализованную структуру данных, обеспечивающую высокую производительность при выполнении любого аналитического запроса, невозможно. Поэтому схема «снежинка» построена таким образом, чтобы обеспечить наивысшую производительность при выполнении одного самого важного запроса, либо группы похожих запросов. В центре схемы «снежинка» располагается таблица фактов (*fact table*), которую окружают таблицами измерений (*dimensional table*). Соединение таблиц измерений с таблицей фактов производится с использованием радиальных связей, в которых таблицы измерений являются родительскими, а таблица фактов — дочерней. Отношение связей — один ко многим.

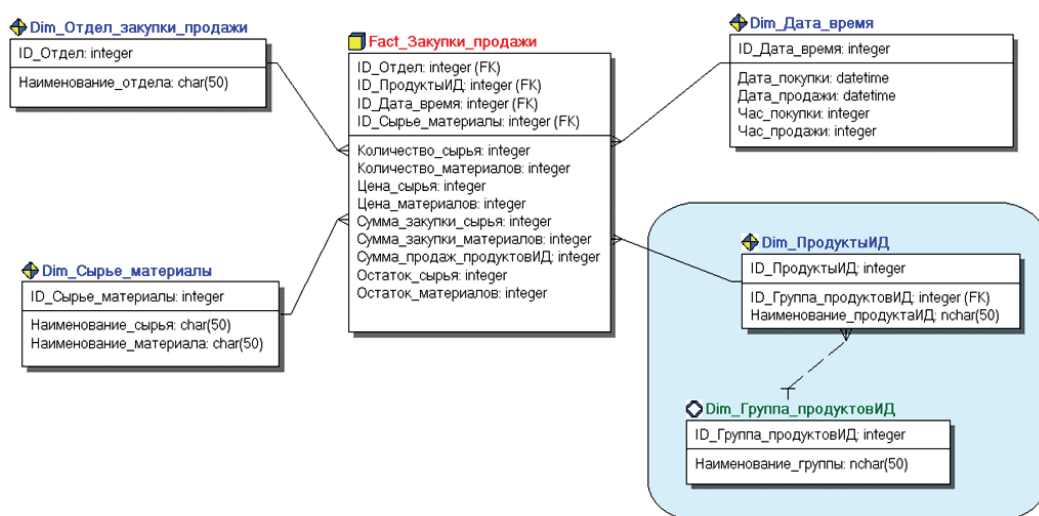


Рис. 11. Размерная модель хранилища данных МИП

Таблица фактов и таблицы измерений связываются идентифицирующими связями, при которых первичные ключи таблиц измерений мигрируют в таблицу фактов в качестве внешних ключей образуя при этом ее составной первичный ключ. Направления связей явно не показываются — они определяются типом таблиц. Составной первичный ключ таблицы фактов целиком состоит из первичных ключей всех таблиц измерений.

Особенностью схемы «снежинка» (в отличие от схемы «звезда» [4]) является обязательное наличие хотя бы одной консольной таблицы (outrigger table), присо-

единяемой к таблице измерений и используемой для ее нормализации. Консольные таблицы могут быть связаны только с таблицами измерений, причем консольная таблица в этой связи родительская, а таблица размерности — дочерняя.

Таким образом, в результате проделанной работы авторами выполнено проектирование как логического, так и физического уровней модели хранилища данных МИП ИТ-профиля, функционирующего при высшем учебном заведении. Применение такой модели позволит обеспечить высокую производительность и упрощение работы с данными о деятельности малого инновационного предприятия.

ЛИТЕРАТУРА

- [1] *Антоников А.А., Чискидов С.В., Павличева Е.Н.* Разработка модуля системы дистанционного обучения для проверки знаний в области программирования // Информационные ресурсы России. — 2012. — № 3. — С. 32—34.
- [2] *Саяпин О.В., Самойлов С.В., Чискидов С.В.* Проектирование АСОИУ: Учеб. пособие. Часть 1: Теоретические основы проектирования автоматизированных систем. — Химки: АГЗ МЧС России, 173 с.
- [3] *Федин Ф.О., Федин Ф.Ф.* Анализ данных. Часть 1: Подготовка данных к анализу: Учеб. пособие. — М.: МГПУ, 2012. — 204 с.
- [4] *Федин Ф.Ф., Павличева Е.Н., Федин Ф.О.* Критерии оценки инновационной деятельности вуза ИТ-профиля // Вестник Российского университета дружбы народов. Серия «Информатизация образования». — 2013. — № 2. — С. 93—99.
- [5] *Федин Ф.О.* Возможности аналитических модулей в информационных системах образовательных учреждений // Вестник Московского городского педагогического университета. Серия «Информатика и информатизация образования». — 2013. — № 1 (25). — С. 119—126.

LITERATURA

- [1] *Antonikov A.A., Chiskidov S.V., Pavlicheva E.N.* Razrabotka modulja sistemy distancionnogo obuchenija dlja proverki znanij v oblasti programmirovanija // Informacionnye resursy Rossii. — 2012. — № 3. — S. 32—34.
- [2] *Sajapin O.V., Samojlov S.V., Chiskidov S.V.* Proektirovanie ASOIU: Ucheb. posobie. Chast' 1: Teoreticheskie osnovy proektirovanija avtomatizirovannyh sistem. — Himki: AGZ MChS Rossii, 173 s.
- [3] *Fedin F.O., Fedin F.F.* Analiz dannyh. Chast' 1: Podgotovka dannyh k analizu: Ucheb. posobie. — M.: MGPU, 2012. — 204 s.
- [4] *Fedin F.F., Pavlicheva E.N., Fedin F.O.* Kriterii ocenki innovacionnoj dejatel'nosti vuza IT-profilja // Vestnik Rossijskogo universiteta druzhby narodov. Serija «Informatizacija obrazovanija». — 2013. — № 2. — S. 93—99.
- [5] *Fedin F.O.* Vozmozhnosti analiticheskikh modulej v informacionnyh sistemah obrazovatel'nyh uchrezhdenij // Vestnik Moskovskogo gorodskogo pedagogicheskogo universiteta. Serija «Informatika i informatizacija obrazovanija». — 2013. — № 1 (25). — S. 119—126.

DEVELOPMENT OF A MODEL OF DATA STORAGE OF INNOVATIVE ENTERPRISE IN A HIGHER EDUCATIONAL INSTITUTION

**F.O. Fedin, S.V. Chiskidov,
E.N. Pavlicheva**

Chair of applied informatics
Moscow city pedagogical university
2-j Tul'skij pereulok, 4, Moscow, Russia, 115191

In work design of logical and physical levels of model of storage of data of the small innovative enterprise of the IT profile functioning at a higher educational institution is executed. Work is performed with use of specialized CASE means of Computer Associates firm — ERWin Process Modeler and ERWin Data Modeler.

Key words: small innovative enterprise, storage of data, Dimensional, SQL inquiry, table of the facts, table of measurements.