

**Министерство науки и высшего образования Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО**

Факультет безопасности информационных технологий

Дисциплина:

«Информационная безопасность баз данных»

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №1

«Инфологическое моделирование баз данных по методу «сущность-связь»»

Выполнил:

Нгуен Хоанг Ву, студент группы N3347



(подпись)

Проверил:

Салихов Максим Русланович, аспирант ФБИТ

(отметка о выполнении)

(подпись)

Санкт-Петербург
2024 г.

СОДЕРЖАНИЕ

Содержание	2
Введение	3
1 Ход Работы	4
1.1 Теоретическая информация	4
1.2 Решение задач	5
1.2.1 Системный анализ информационной системы	5
1.2.2 Выделение сущностей и построение ER-диаграммы	6
1.2.3 Преобразование ER-диаграммы в схему отношений с помощью правил формирования предварительных отношений.	8
1.2.4 Приведение схемы предварительных отношений к 3НФ.	9
1.2.5 Моделирование уровня представлений ИС супермаркета	11
Вывод	13

ВВЕДЕНИЕ

Цель работы - Изучение способов семантического представления баз данных (БД), получение навыков инфологического проектирования с использованием нотации «сущность-связь».

Задание:

- самостоятельно выбрать информационную систему (ИС), для которой будет составлена база данных. Выбранные информационные системы у студентов не должны повторяться. Для всего потока будет создана таблица, в которую можно будет внести выбранную информационную систему;
- используя метод «сущность-связь», провести моделирование базы данных для выбранной в пункте 1 информационной системы.

1 ХОД РАБОТЫ

1.1 Теоретическая информация

- 1 этап моделирования. Системный анализ информационной системы. На данном этапе подробно описывается:
 - для каких именно процессов в выбранной ИС разрабатывается база данных. Какие задачи планируется решать с помощью разрабатываемой базы данных.
 - что является источником данных, откуда берутся данные, в каком формате/виде поступает информация в БД, с какой частотой обновляется;
 - кто (какие классы пользователей) или что (ИС более высокого порядка) является потребителем информации. В каком формате/виде представляется информация для различных классов потребителей. Выделите как минимум 2 класса потребителя информации, не считая администратора БД.
 - описание ограничений на сущности и связи в будущей модели базы данных. Например, если разрабатывается база для ИС университета. На данном этапе могут быть описаны следующие очевидные ограничения (в одном учебном потоке не может быть одновременно занятий по разным дисциплинам; один преподаватель не может в одно и тоже время проводить занятия по различным дисциплинам и пр.)
- 2 этап моделирования. Выделение сущностей и построение ER-диаграмм. На данном этапе формируется список сущностей для разрабатываемой БД. Кратко описывается, что моделирует и отражает каждая сущность в рамках вашей ИС. Указывается, что представляет собой в системе каждый экземпляр сущности. Далее проводится анализ связей между сущностями с указанием типа связи и класса принадлежности. В итоге данного этапа должна получиться ER-диаграмма. Минимальное количество отношений, которое должно получиться после моделирования – 7. Минимальное количество атрибутов (столбцы) в отношениях - 3. Отношения-связи между таблицами не считаются при подсчете минимально необходимого количества отношений
- 3 этап моделирования. Преобразование ER-диаграммы в схему отношений с помощью правил формирования предварительных отношений;
- 4 этап моделирования. Приведение отношений БД к 3НФ.
- 5 этап моделирования. Создайте по крайней мере 2 представления (view) на каждого потребителя информации из вашей БД. Представление должно основываться на запросе, который затрагивает атрибуты из 2-х и более связанных отношений.

1.2 Решение задач

1.2.1 Системный анализ информационной системы

В качестве информационной системы, для которой будет составлена БД, была выбрана «Информационная система супермаркета». База данных разрабатывается для управления запасами, продажами и информацией о клиентах супермаркета.

Задачи, которые будут решаться с помощью разрабатываемой базы данных, включают:

- отслеживание уровня запасов;
- управление транзакциями продаж;
- анализ моделей покупок клиентов.

Источниками данных для базы данных супермаркета являются:

- POS-системы;
- системы управления запасами;
- опросы клиентов.

Данные поступают в электронном формате (например, CSV, Excel) и обновляются в режиме реального времени для POS-систем и систем управления запасами, в то время как опросы клиентов обновляются периодически (например, ежемесячно, ежеквартально).

Потребителями из разрабатываемой БД являются:

- менеджеры супермаркетов;
- руководители отделов;
- аналитики по маркетингу.

Информация представлена в различных форматах, включая:

- графические отчеты;
- табличные данные;
- диаграммы.

Например, менеджеры супермаркетов могут просматривать уровни запасов в табличном формате, в то время как аналитики по маркетингу могут просматривать модели покупок клиентов в графическом формате. Два класса потребителей информации — менеджеры супермаркетов и аналитики по маркетингу.

Описание ограничений на сущности моделируемой базы данных.

Некоторые ограничения на сущности и отношения в модели базы данных супермаркета включают:

Каждый продукт может иметь только одну категорию, но может принадлежать к нескольким подкатегориям.

Каждый клиент может иметь только один основной адрес, но может иметь несколько вторичных адресов.

Каждая транзакция продажи может иметь только одного клиента, но может включать несколько продуктов.

Каждый продукт может иметь только одного поставщика, но может иметь несколько поставщиков.

1.2.2 Выделение сущностей и построение ER-диаграммы.

На основе проведенного анализа ИС, для которой разрабатывается БД, можно выделить следующие сущности:

- «Продукт»
- «Клиент»
- «Продажи_Транзакция»

Сущность «Продукт» характеризуется атрибутами «Идентификатор», «Цена», «Название» и «Описание». Ключом этой сущности является атрибут «Идентификатор», так как каждый продукт имеет свой уникальный идентификатор.

Сущность «Клиент» представляет покупателя, который совершает покупку в супермаркете. Каждый клиент имеет уникальное лицо с собственной информацией, такой как имя, адрес и номер телефона. Ключом этой сущности является атрибут «Идентификатор для каждого покупателя»

Сущность «Продажи_Транзакция» представляет транзакцию продажи между покупателем и супермаркетом. Каждый экземпляр транзакции продажи представляет собой уникальную покупку, совершенную клиентом, включая дату, сумму транзакции. Ключом этой сущности является атрибут «Идентификатор для каждой транзакции продажи»

Связи:

Клиент может совершать множество транзакций продаж (один ко многим).

Клиент -> Транзакция_продаж (идентификатор_клиента)

Продукт может участвовать во множестве транзакций продаж (один ко многим).

Продукт -> Транзакция_продаж (идентификатор_продукта)

Транзакция продаж связана с одним клиентом (многие к одному).

Транзакция_продаж -> Клиент (идентификатор_клиента)

Транзакция продаж связана с одним продуктом (многие к одному).

Транзакция_продаж -> Продукт (идентификатор_продукта)

Клиент может приобретать множество продуктов (многие ко многим).

Клиент -> Продукт (через Транзакцию_продаж)

Продукт может приобретаться многими клиентами (многие ко многим).

Продукт -> Клиент (через Транзакцию_продаж)

В соответствии с выделенными сущностями и связями между ними ER-диаграмма для рассматриваемой предметной области будет выглядеть так, как показано на рисунке 1.

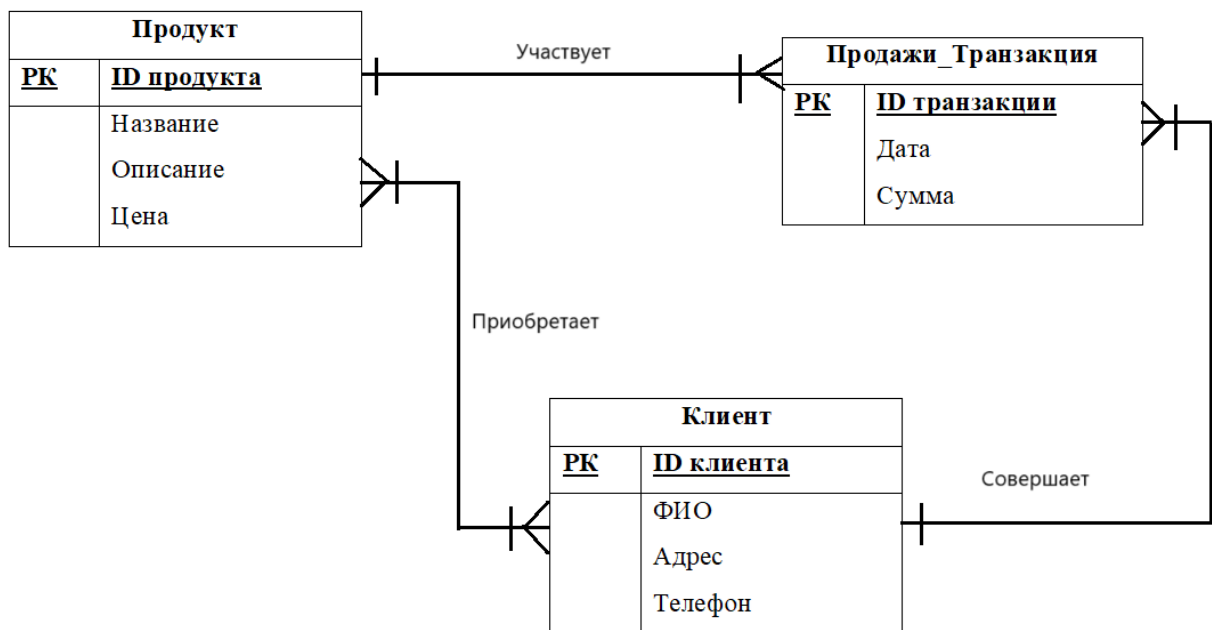


Рисунок 1 – ER-диаграмма разрабатываемой БД

1.2.3 Преобразование ER-диаграммы в схему отношений с помощью правил формирования предварительных отношений.

Реляционная диаграмма состоит из трех таблиц: Продукт, Клиент и Продаж_Транзакция. Таблица Продукт содержит столбцы для ID продукта (первичный ключ), название, цена и описание. Таблица Клиент содержит столбцы для ID клиента (первичный ключ), ФИО, Адрес и Телефон. Таблица Продаж_Транзакция содержит столбцы для ID транзакции (первичный ключ), Дата, Сумма, ID Клиента (внешний ключ, ссылающийся на Client.Id) и ID продукта (внешний ключ, ссылающийся на Product.Id).

Связи между таблицами следующие: клиент может иметь несколько транзакций продаж (один ко многим), продукт может быть вовлечен в несколько транзакций продаж (один ко многим), транзакция продаж связана с одним клиентом (многие к одному), а транзакция продаж связана с одним продуктом (многие к одному). Кроме того, отношения «многие ко многим» разрешаются с помощью таблицы Продаж_Транзакция, которая действует как мост между таблицами Клиент и Продукт, позволяя клиенту приобретать несколько продуктов, а продукт — приобретать многим клиентам.

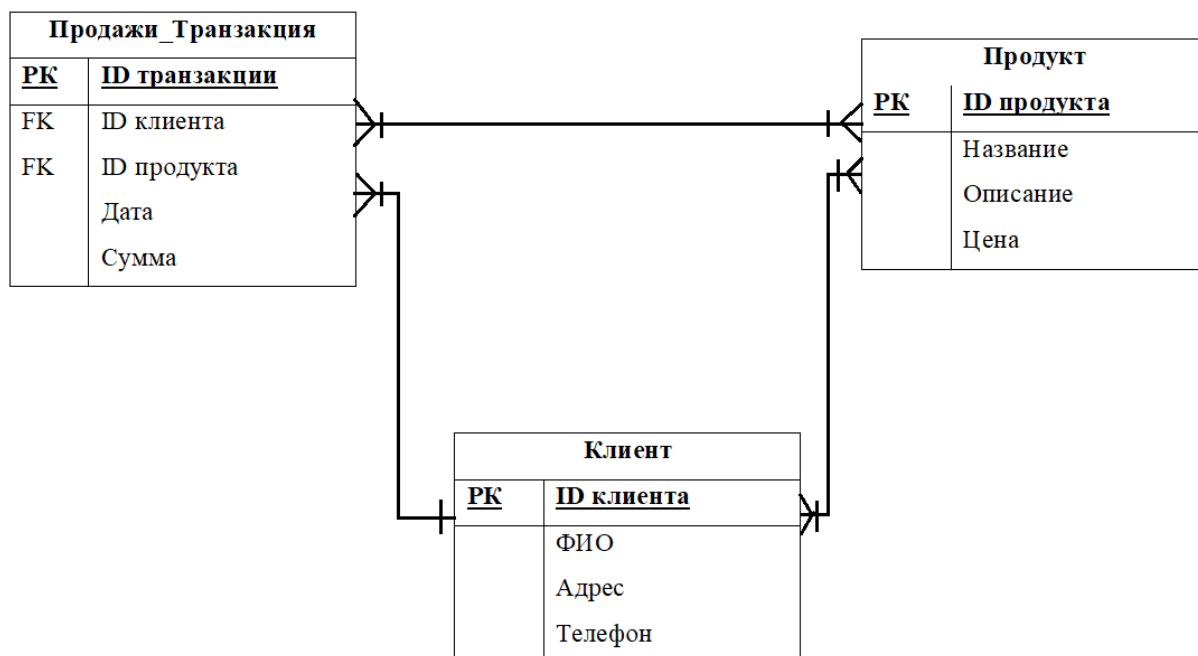


Рисунок 2 – Итоговая схема предварительных отношений

1.2.4 Приведение схемы предварительных отношений к ЗНФ.

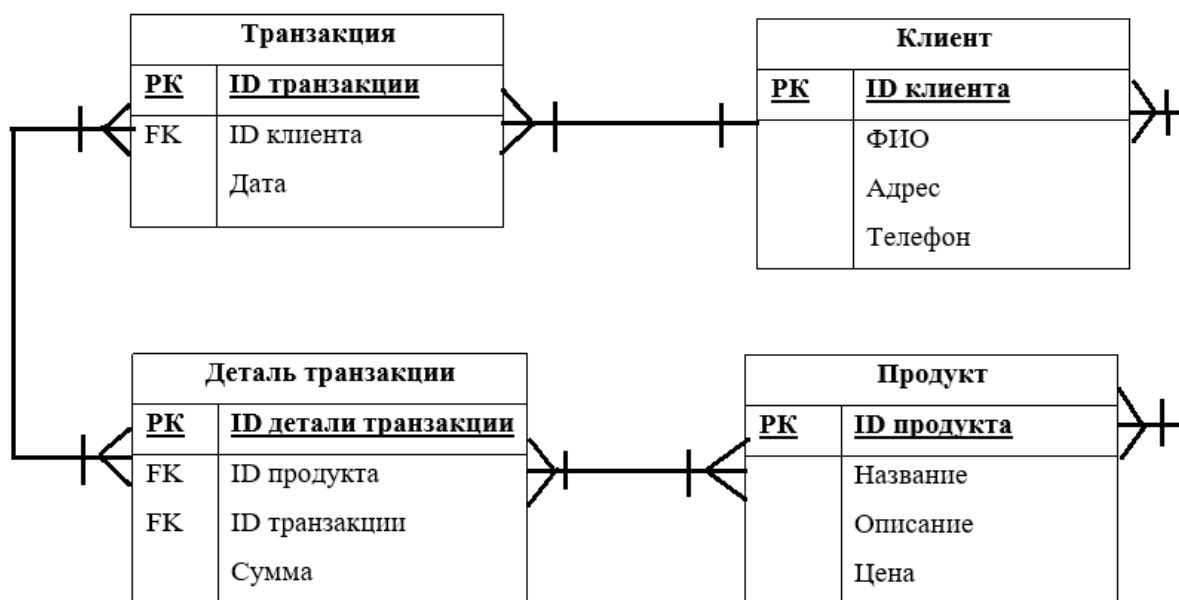


Рисунок 3 – Итоговая схема предварительных отношений к ЗНФ

Устранение повторяющихся групп

Исходная схема отношений имеет повторяющуюся группу в отношении «Продаж_Транзакция», где для каждой транзакции хранятся несколько значений product_id и transaction_amount. Чтобы устранить это, мы разбиваем отношение «Продаж_Транзакция» на два отношения: «Транзакция» и «Деталь транзакции». Отношение «Деталь транзакции» хранит каждый продукт, купленный в транзакции, вместе с соответствующей суммой транзакции.

Устранение частичных зависимостей

В отношении «Деталь транзакции» атрибут transaction_amount частично зависит от атрибута product_id. Чтобы устранить это, мы делаем transaction_amount полностью зависимым от составного ключа transaction_id и product_id в отношении «Деталь транзакции».

Устранение транзитивных зависимостей

В схеме отношений нет транзитивных зависимостей.

Результирующая схема отношений в 3НФ

Результирующая схема отношений в 3НФ состоит из следующих отношений:

Отношение клиента: customer_id, customer_name, customer_address, customer_phone

Отношение транзакции: transaction_id, customer_id, transaction_date

Отношение детали транзакции: transaction_id, product_id, transaction_amount

Отношение продукта: product_id, product_name

Это сокращение до 3НФ устраняет все повторяющиеся группы, частичные зависимости и транзитивные зависимости, что приводит к более нормализованной и эффективной схеме отношений.

1.2.5 Моделирование уровня представлений ИС супермаркета

Выделим три обязательных потребителя информации из моей базы данных. Первый потребитель – это представитель службы поддержки клиентов . Второй потребитель – это менеджер по продукту. Третий потребитель - это маркетинговая команда.

1.2.5.1 Потребитель: Представитель службы поддержки клиентов

Представление 1: История покупок клиентов

Список атрибутов, отображаемых в рамках представления «История покупок клиентов»:

- Идентификатор клиента (Отношение «Клиент»)
- Идентификатор транзакции (Отношение «Продаж_Транзакция»)
- Дата транзакции (Отношение «Продаж_Транзакция»)
- Сумма транзакции (Отношение «Продаж_Транзакция»)

Представление 2: Общие расходы клиента

Список атрибутов, отображаемых в рамках представления «Общие расходы клиента»:

- Идентификатор клиента (Отношение «Клиент»)
- Дата транзакции (Отношение «Продаж_Транзакция»)
- Сумма транзакции (Отношение «Продаж_Транзакция»)

1.2.5.2 Потребитель: Менеджер по продукту

Представление 1: Продаж продукции по дате

Список атрибутов, отображаемых в рамках представления «Продаж продукции по дате»:

- Идентификатор продукта (Отношение «Продукт»)
- Дата транзакции (Отношение «Продаж_Транзакция»)

Представление 2: Бестселлеры продукции

Список атрибутов, отображаемых в рамках представления «Бестселлеры продукции»:

- Идентификатор продукта (Отношение «Продукт»)

1.2.5.3 Потребитель: Маркетинговая команда

Представление 1 :Демография клиентов

Список атрибутов, отображаемых в рамках представления «Демография клиентов»:

- Идентификатор клиента (Отношение «Клиент»)
- Адрес (Отношение «Клиент»)
- ФИО (Отношение «Клиент»)
- Телефон (Отношение «Клиент»)

Представление 2:Популярность продуктов по регионам

Список атрибутов, отображаемых в рамках представления «Популярность продуктов по регионам»:

- Идентификатор продукта (Отношение «Продукт»)
- Адрес (Отношение «Клиент»)

ВЫВОД

В заключение, моя задача успешно продемонстрировала применение методов семантического представления в базах данных, в частности, используя нотацию «сущность-связь» для инфологического проектирования. Нормализуя схему базы данных до 3НФ, я эффективно представил данные семантически, выделив связи между четырьмя ключевыми сущностями: Клиент, Транзакция, Деталь транзакции и Продукт. Это упражнение усилило важность семантического представления для обеспечения согласованности данных, снижения избыточности и улучшения целостности данных в проектировании базы данных.