

ИНТЕГРИРОВАННЫЕ СИСТЕМЫ УПРАВЛЕНИЯ

Лекция №8

**Классификация и кодирование
информации**

Система классификации

Важным понятием при работе с информацией является *классификация объектов*.

Классификация - система распределения объектов (предметов, явлений, процессов, понятий) по классам в соответствии с определенным признаком.

Под **объектом** понимается любой предмет, процесс, явление материального или нематериального свойства. Система классификации позволяет сгруппировать объекты и выделить определенные классы, которые будут характеризоваться рядом общих свойств. Классификация объектов – это процедура группировки на качественном уровне, направленная на выделение однородных свойств. Применительно к информации как к объекту классификации выделенные классы называют **информационными объектами**.

Система классификации

Свойства информационного объекта определяются информационными параметрами, называемыми *реквизитами*. Реквизиты представляются либо числовыми данными, например вес, стоимость, год, либо признаками, например цвет, марка машины, фамилия.

Реквизит - логически неделимый информационный элемент, описывающий определенное свойство объекта, процесса, явления и т.п.

Кроме выявления общих свойств информационного объекта классификация нужна для разработки правил (алгоритмов) и процедур обработки информации, представленной совокупностью реквизитов.

Система классификации

При любой классификации желательно, чтобы соблюдались следующие требования:

- полнота охвата объектов рассматриваемой области;
- однозначность реквизитов;
- возможность включения новых объектов.

В любой стране разработаны и применяются государственные, отраслевые, региональные классификаторы. Например, классифицированы: отрасли промышленности, оборудование, профессии, единицы измерения, статьи затрат и т.д.

Классификатор - систематизированный свод наименований и кодов классификационных группировок.

Система классификации

При классификации широко используются понятия **классификационный признак** и *значение классификационного признака*, которые позволяют установить сходство или различие объектов. Возможен подход к классификации с объединением этих двух понятий в одно, названное как признак классификации. Признак классификации имеет также синоним основание деления.

Разработаны три метода классификации объектов:

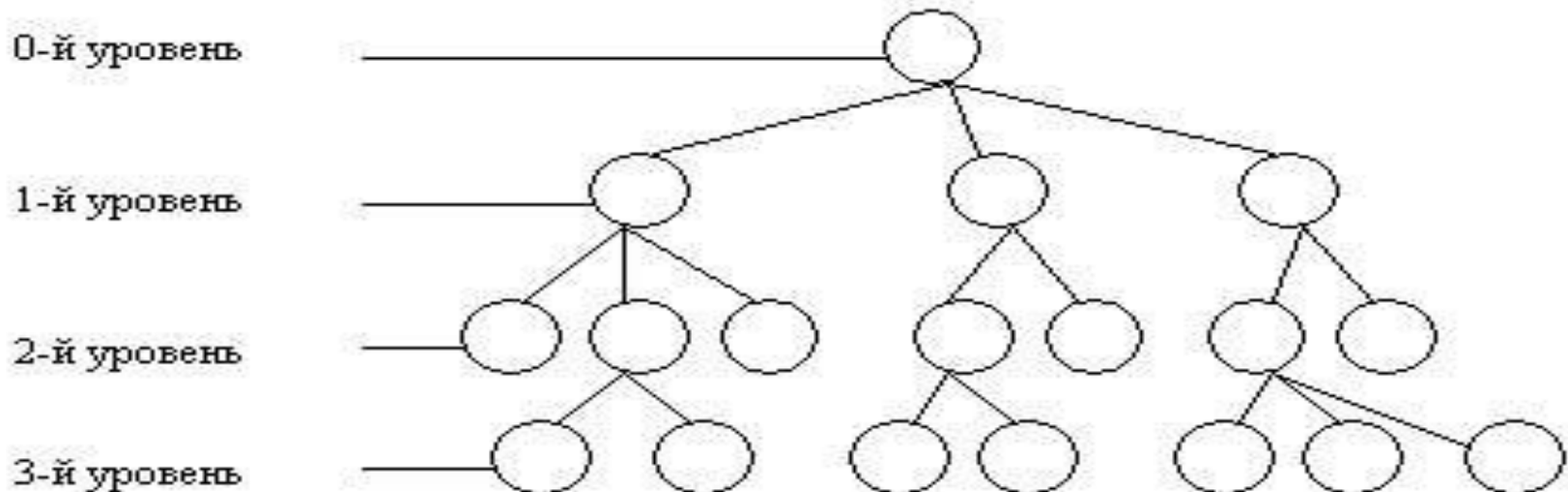
- иерархический,
- фасетный,
- дескрипторный.

Эти методы различаются разной стратегией применения классификационных признаков. Рассмотрим основные идеи этих методов для создания систем классификации.

Иерархическая система классификации

Иерархическая система классификации строится следующим образом:

- исходное множество элементов составляет 0-й уровень и делится в зависимости от выбранного классификационного признака на классы (группировки), которые образуют 1-й уровень;
- каждый класс 1-го уровня в соответствии со своим, характерным для него классификационным признаком делится на подклассы, которые образуют 2-й уровень;
- каждый класс 2-го уровня аналогично делится на группы, которые образуют 3-й уровень и т.д.



Иерархическая система классификации

Учитывая достаточно жесткую процедуру построения структуры классификации, необходимо перед началом работы определить ее цель, т.е. какими свойствами должны обладать объединяемые в классы объекты. Эти свойства принимаются в дальнейшем за признаки классификации.

Запомните! В иерархической системе классификации из-за жесткой структуры особое внимание следует уделить выбору классификационных признаков.

В иерархической системе классификации каждый объект на любом уровне должен быть отнесен к одному классу, который характеризуется конкретным значением выбранного классификационного признака. Для последующей группировки в каждом новом классе необходимо задать свои классификационные признаки и их значения. Таким образом, выбор классификационных признаков будет зависеть от семантического содержания того класса, для которого необходима группировка на последующем уровне иерархии.

Иерархическая система классификации

Количество уровней классификации, соответствующее числу признаков, выбранных в качестве основания деления, характеризует ***глубину классификации***.

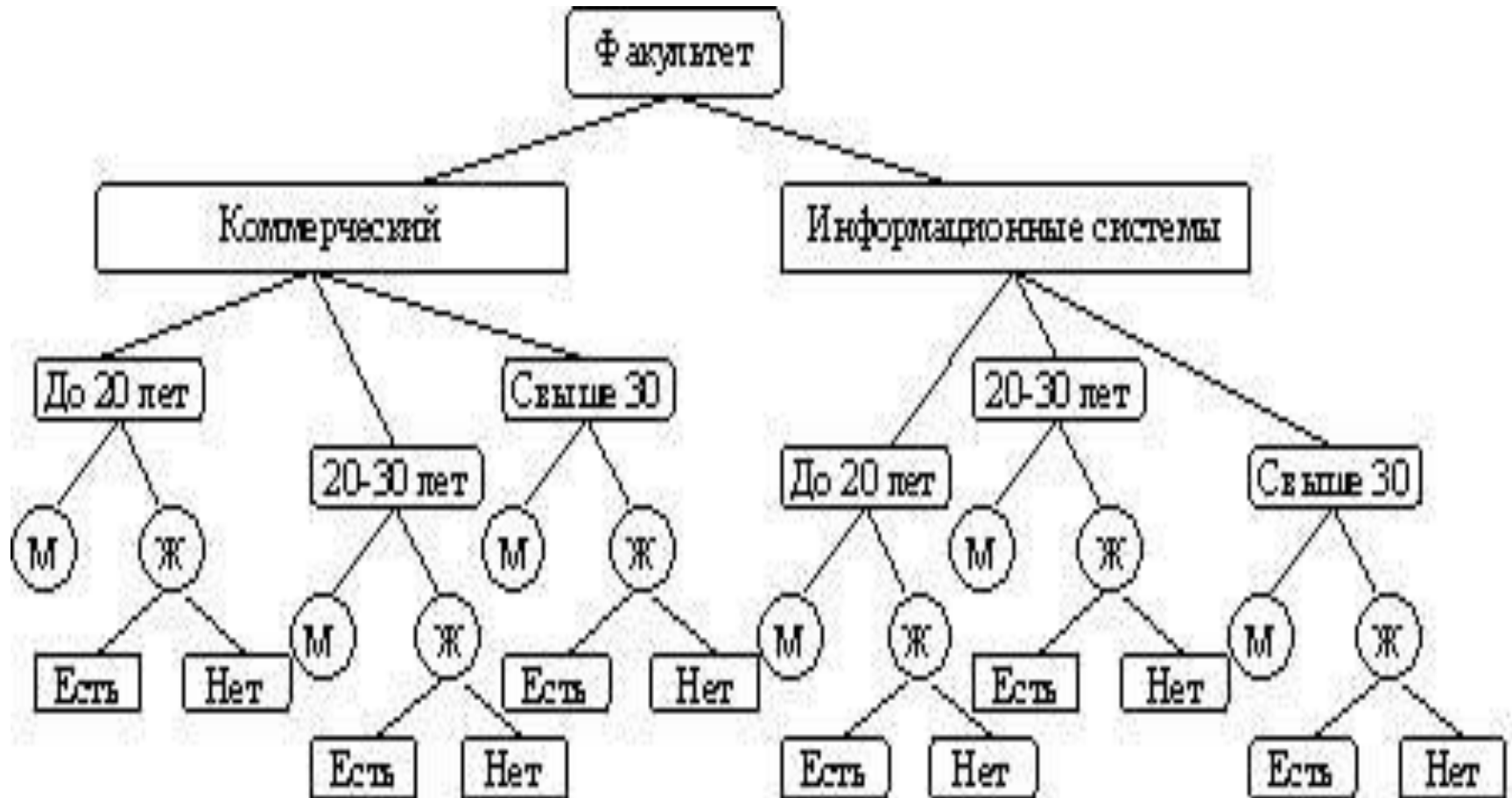
Достоинства иерархической системы классификации:

- простота построения;
- использование независимых классификационных признаков в различных ветвях иерархической структуры.

Недостатки иерархической системы классификации:

- жесткая структура, которая приводит к сложности внесения изменений, так как приходится перераспределять все классификационные группировки;
- невозможность группировать объекты по заранее не предусмотренным сочетаниям признаков.

Иерархическая система классификации



Пример иерархической системы классификации для информационного объекта «Факультет»

Фасетная система классификации

Фасетная система классификации в отличие от иерархической позволяет выбирать признаки классификации независимо как друг от друга, так и от семантического содержания классифицируемого объекта. Признаки классификации называются **фасетами** (facet - рамка). Каждый фасет (Φ_i) содержит совокупность однородных значений данного классификационного признака. Причем значения в фасете могут располагаться в произвольном порядке хотя предпочтительнее их упорядочение.

Фасетная система классификации

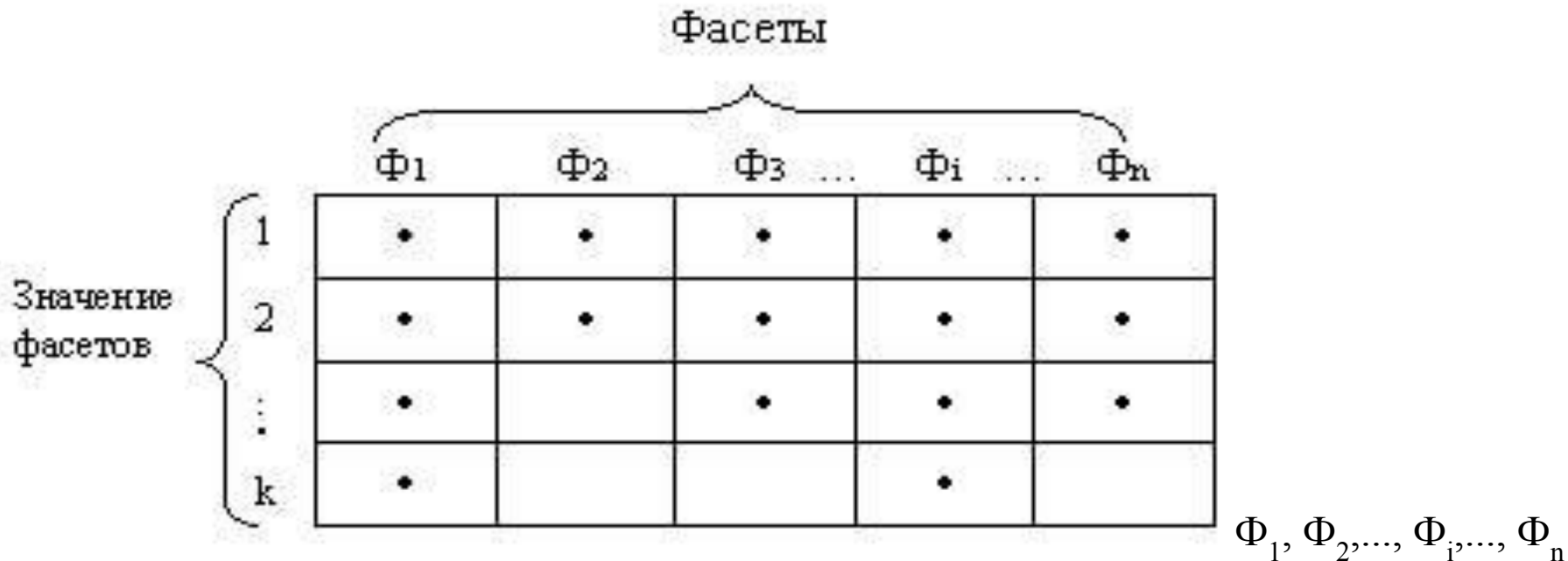


Схема построения фасетной системы классификации в виде таблицы отображена на рисунке. Названия столбцов соответствуют выделенным классификационным признакам (фасетам), обозначенным $\Phi_1, \Phi_2, \dots, \Phi_i, \dots, \Phi_n$. Например, цвет, размер одежды, вес и т.д. Произведена нумерация строк таблицы. В каждой клетке таблицы хранится конкретное **значение фасета**. Например, фасет *цвет*, обозначенный Φ_2 , содержит значения: красный, белый, зеленый, черный, желтый.

Фасетная система классификации

Процедура классификации состоит в присвоении каждому объекту соответствующих значений из фасетов. При этом могут быть использованы не все фасеты. Для каждого объекта задается конкретная группировка фасетов структурной формулой, в которой отражается их порядок следования:

$$K_s = (\Phi_1, \Phi_2, \dots, \Phi_i, \dots, \Phi_n),$$

где Φ_i - i -й фасет; n - количество фасетов.

При построении фасетной системы классификации необходимо, чтобы значения, используемые в различных фасетах, не повторялись. Фасетную систему легко можно модифицировать, внося изменения в конкретные значения любого фасета.

Фасетная система классификации

Достоинства фасетной системы классификации:

- возможность создания большой емкости классификации, т.е. использования большого числа признаков классификации и их значений для создания группировок;
- возможность простой модификации всей системы классификации без изменения структуры существующих группировок.

Недостатком фасетной системы классификации является сложность ее построения, так как необходимо учитывать все многообразие классификационных признаков.

Дескрипторная система классификации

Для организации поиска информации, для ведения тезаурусов (словарей) эффективно используется дескрипторная (описательная) система классификации, язык которой приближается к естественному языку описания информационных объектов. Особенно широко она используется в библиотечной системе поиска.

Суть дескрипторного метода классификации заключается в следующем:

- отбирается совокупность ключевых слов или словосочетаний, описывающих определенную предметную область или совокупность однородных объектов. Причем среди ключевых слов могут находиться синонимы;
- выбранные ключевые слова и словосочетания подвергаются *нормализации*, т.е. из совокупности синонимов выбирается один или несколько наиболее употребимых;
- создается **словарь дескрипторов**, т.е. словарь ключевых слов и словосочетаний, отобранных в результате процедуры нормализации.

Дескрипторная система классификации

Между дескрипторами устанавливаются связи, которые позволяют расширить область поиска информации. Связи могут быть трех видов:

- **синонимические** указывающие некоторую совокупность ключевых слов как синонимы;
- **родо-видовые**, отражающие включение некоторого класса объектов в более представительный класс;
- **ассоциативные**, соединяющие дескрипторы, обладающие общими свойствами.

Система кодирования

Система кодирования применяется для замены названия объекта на условное обозначение (код) в целях обеспечения удобной и более эффективной обработки информации.

Система кодирования - совокупность правил кодового обозначения объектов.

Код строится на базе алфавита, состоящего из букв, цифр и других символов. Код характеризуется:

- **длиной** - число позиций в коде;
- **структурой** - порядок расположения в коде символов, используемых для обозначения классификационного признака.

Процедура присвоения объекту кодового обозначения называется **кодированием**. Можно выделить две группы методов, используемых в системе кодирования, которые образуют:

- **классификационную систему кодирования**, ориентированную на проведение предварительной классификации объектов либо на основе иерархической системы, либо на основе фасетной системы;
- **регистрационную систему кодирования**, не требующую предварительной классификации объектов.

Система кодирования



Система кодирования, использующая разные методы

Классификационное кодирование

Классификационное кодирование применяется после проведения классификации объектов. Различают последовательное и параллельное кодирование.

Последовательное кодирование используется для иерархической классификационной структуры. Суть метода заключается в следующем: сначала записывается код старшей группировки 1-го уровня, затем код группировки 2-го уровня, затем код группировки 3-го уровня и т.д. В результате получается кодовая комбинация, каждый разряд которой содержит информацию о специфике выделенной группы на каждом уровне иерархической структуры. Последовательная система кодирования обладает теми же достоинствами и недостатками, что и иерархическая система классификации.

Параллельное кодирование используется для фасетной системы классификации. Суть метода заключается в следующем: все фасеты кодируются независимо друг от друга; для значений каждого фасета выделяется определенное количество разрядов кода. Параллельная система кодирования обладает теми же достоинствами и недостатками, что и фасетная система классификации.

Регистрационное кодирование

Регистрационное кодирование используется для однозначной идентификации объектов и не требует предварительной классификации объектов. Различают порядковую и серийно-порядковую систему.

Порядковая система кодирования предполагает последовательную нумерацию объектов числами натурального ряда. Этот порядок может быть случайным или определяться после предварительного упорядочения объектов, например по алфавиту. Этот метод применяется в том случае, когда количество объектов невелико, например кодирование названий факультетов университета, кодирование студентов в учебной группе.

Серийно-порядковая система кодирования предусматривает предварительное выделение групп объектов, которые составляют серию, а затем в каждой серии производится порядковая нумерация объектов. Каждая серия также будет иметь порядковую нумерацию. По своей сути серийно-порядковая система является смешанной: классифицирующей и идентифицирующей. Применяется тогда, когда количество групп невелико.

Классификация информации по разным признакам

Любая классификация всегда относительна. Один и тот же объект может быть классифицирован по разным признакам или критериям. Часто встречаются ситуации, когда в зависимости от условий внешней среды объект может быть отнесен к разным классификационным группировкам. Эти рассуждения особенно актуальны при классификации видов информации без учета ее предметной ориентации, так как она часто может быть использована в разных условиях, разными потребителями, для разных целей.

Классификация информации по разным признакам

На рисунке приведена одна из схем классификации циркулирующей в организации (фирме) информации. В основу классификации положено пять наиболее общих признаков: место возникновения, стадия обработки, способ отображения, стабильность, функция управления.



Классификация информации по разным признакам

Место возникновения. По этому признаку информацию можно разделить на входную, выходную, внутреннюю, внешнюю.

Входная информация - это информация, поступающая в фирму или ее подразделения.

Выходная информация - это информация, поступающая из фирмы в другую фирму, организацию (подразделение).

Одна и та же информация может являться входной для одной фирмы, а для другой, ее вырабатывающей, выходной. По отношению к объекту управления (фирма или ее подразделения: цех, отдел, лаборатория) информация может быть определена как внутренняя, так и внешняя.

Внутренняя информация возникает внутри объекта, **внешняя** информация - за пределами объекта.

Классификация информации по разным признакам

Стадия обработки. По стадии обработки информация может быть первичной, вторичной, промежуточной, результатной.

Первичная информация - это информация, которая возникает непосредственно в процессе деятельности объекта и регистрируется на начальной стадии.

Вторичная информация - это информация, которая получается в результате обработки первичной информации и может быть промежуточной и результатной.

Промежуточная информация используется в качестве исходных данных для последующих расчетов.

Результатная информация получается в процессе обработки первичной и промежуточной информации и используется для выработки управленческих решений.

Классификация информации по разным признакам

Способ отображения. По способу отображения информация подразделяется на текстовую и графическую.

Текстовая информация - это совокупность алфавитных, цифровых и специальных символов, с помощью которых представляется информация на физическом носителе (бумага, изображение на экране дисплея).

Графическая информация - это различного рода графики, диаграммы, схемы, рисунки и т.д.

Классификация информации по разным признакам

Стабильность. По стабильности информация может быть переменной (текущей) и постоянной (условно-постоянной).

Переменная информация отражает фактические количественные и качественные характеристики производственно-хозяйственной деятельности фирмы. Она может меняться для каждого случая как по назначению, так и по количеству. *Например, количество произведенной продукции за смену, еженедельные затраты на доставку сырья, количество исправных станков и т.п.*

Постоянная (условно-постоянная) информация - это неизменная и многократно используемая в течение длительного периода времени информация. Постоянная информация может быть справочной, нормативной, плановой:

- постоянная справочная информация включает описание постоянных свойств объекта в виде устойчивых длительное время признаков. *Например, табельный номер служащего, профессия работника, номер цеха и т.п.;*
- постоянная нормативная информация содержит местные, отраслевые и общегосударственные нормативы. *Например, размер налога на прибыль, стандарт на качество продуктов определенного вида, размер минимальной оплаты труда, тарифная сетка оплаты государственным служащим;*
- постоянная плановая информация содержит многократно используемые в фирме плановые показатели. *Например, план выпуска телевизоров, план подготовки специалистов определенной квалификации.*

Классификация информации по разным признакам

Функция управления. По функциям управления обычно классифицируют **экономическую информацию**. При этом выделяют следующие группы: плановую, нормативно-справочную, учетную и оперативную (текущую).

Плановая информация - информация о параметрах объекта управления на будущий период. На эту информацию идет ориентация всей деятельности фирмы.

Нормативно-справочная информация содержит различные нормативные и справочные данные. Ее обновление происходит достаточно редко.

Учетная информация - это информация, которая характеризует деятельность фирмы за определенный прошлый период времени. На основании этой информации могут быть проведены следующие действия: скорректирована плановая информация, сделан анализ хозяйственной деятельности фирмы, приняты решения по более эффективному управлению работами и пр.

Оперативная (текущая) информация - это информация, используемая в оперативном управлении и характеризующая производственные процессы в текущий (данный) период времени. К оперативной информации предъявляются серьезные требования по скорости поступления и обработки, а также по степени ее достоверности. От того, насколько быстро и качественно проводится ее обработка, во многом зависит успех фирмы на рынке.