

Тема 6. Методы классификации.**2 часа**

Множество любых объектов: потребительских товаров или промышленных изделий, растений или бабочек, форм документов или программных средств – легче поддается изучению, если оно упорядочено, систематизировано.

Классификация – это упорядочение, «разделение множества объектов на подмножества по их сходству или различию в соответствии с принятыми методами»¹. Классификация фиксирует закономерные связи между классами объектов с целью определения места объекта в системе, которое указывает на его свойства. 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.

Классификация носит всеобщий характер вследствие той роли, которую она играет как инструмент научного познания, прогнозирования и управления. Одновременно классификация выполняет функцию объективного отражения и фиксации результатов этого познания. При этом характер классификационной схемы, состав признаков классификации и глубина классификации определяется целями классификации, типом объектов классификации, условиями, в которых классификация будет использоваться.

При классификации технико-экономической и социальной информации используются иерархический и фасетный методы. Для изучения методов классификации введем термины (по ПР 50.1.19-2000):

Объект классификации: Элемент классифицируемого множества.

¹ ПР 50.1.019-2000. Правила по стандартизации. Основные положения ЕСКК ТЭСИ и УСД в РФ.

Признак классификации: Свойство или характеристика объекта классификации, по которому проводится классификация.

Классификационная группировка: Подмножество объектов, полученное в результате классификации.

Степень классификации: Этап классификации при иерархическом методе классификации, результатом которого является совокупность классификационных группировок.

Глубина классификации: Число ступеней классификации.

Методы классификации должны отвечать следующим общим требованиям:

- обладать достаточной емкостью;
- иметь достаточную и экономически оправданную глубину;
- обладать определенной гибкостью и избыточностью для возможности расширения множества классифицируемых объектов, группировок и признаков и внесения необходимых изменений без нарушения структуры классификации;
- учитывать необходимость сопряжения с другими классификациями однородных объектов;
- обеспечивать наибольшую эффективность обработки информации средствами вычислительной техники;
- предоставлять возможность ведения создаваемого классификатора как в ручном, так и автоматизированном варианте.

Иерархический метод классификации обеспечивает последовательное разделение множества объектов на соподчиненные группировки. Исходное множество объектов классификации делится сначала по какому-либо признаку на крупные группировки, каждая из которых, в свою очередь дробится на ряд последующих группировок, конкретизируя свойства объектов в виде иерархического логического дерева.



Рис. 1. Пример иерархического метода классификации.

Таким образом, под иерархическим методом классификации понимается такой метод, при котором заданное множество объектов последовательно делится на подчиненные подмножества. При построении классификаторов иерархическим методом классификации необходимо соблюдать следующие правила:

- деление каждой классификационной группировки должно производиться только по одному основанию;
- получаемые в результате деления группировки не должны пересекаться, то есть содержать аналогичной информации, и должны относиться только к одной вышестоящей группировке;
- деление исходного множества на подмножества должно быть последовательным, без пропусков и без добавления промежуточного уровня классификации;
- классифицирование должно производиться таким образом, чтобы сумма подмножеств деления составляла делимое множество.

Иерархический метод классификации характеризуется глубиной классификации и емкостью. Количество ступеней определяет глубину классификации, которая устанавливается в зависимости от степени конкретизации группировок и числа признаков, необходимых для решения конкретных задач. От глубины и количества группировок, образуемых на каждой ступени классификации, зависит емкость системы. Как правило, наибольшее количество последующих группировок устанавливается постоянным, либо для всей классификации, либо

для данной ступени. Для систем классификации технико-экономической информации это число всегда равно десяти или кратному десяти. Например, в Общероссийском классификаторе управленческой документации (ОКУД) принят иерархический метод классификации с тремя ступенями: класс, подкласс и регистрационный номер.

Преимущества иерархического метода:

1) При иерархическом методе классификации практически не ограничивается глубина классификации информации, что дает возможность более детально анализировать предметы, явления или документы.

2) Большая информационная емкость иерархического метода классификации позволяет использовать его для кодирования больших объемов технико-экономической информации.

Недостатки иерархического метода:

1) Недостаточная гибкость структуры, обусловленная фиксированностью признаков классификации и заранее установленным порядком их следования, не допускающим включения новых объектов и классификационных группировок. Вследствие этого, изменение любого признака ведет к перераспределению классификационных группировок и необходимости переработки классификатора. Поэтому в классификаторах, построенных на основе иерархического метода, должны предусматриваться значительные резервные емкости.

2) Этот метод классификации не позволяет осуществлять информационный поиск по любому произвольному сочетанию признаков.

Под **фасетным методом** классификации понимается такой метод, при котором классифицируемое множество делится на группировки независимо, по различным признакам классификации. Он не имеет жесткой классификационной структуры и заранее построенных конечных группировок. Классификационные группировки при фасетном методе классификации образуются путем комбинаций значений признаков, взятых из соответствующих фасетов. Т.е. фасет – это признак классификации со своим набором значений.

В качестве примера вкратце рассмотрим Отраслевой классификатор банковской документации. В основе Классификатора банка лежит фасетный метод классификации. В качестве классификационных группировок выделены реквизиты документов, используемых в документообороте Банка. Фасеты классифицируемого множества располагаются в Классификаторе в виде простого перечисления и кодируются порядковыми номерами. Внутри каждого фасета значения фасета располагаются в виде простого или алфавитного перечисления и также кодируются порядковыми номерами. Количество фасетов и количество значений фасетов неограниченно, должно быть достаточно для полной идентификации документа и наращивается в процессе ведения классификатора. Идентификация документа при его регистрации и дальнейшей работы с документом производится путем выборки нужных фасетов и значений фасетов. Последовательность расположения фасетов и их значений при кодировании документа определяется исполнителем и зависит от характера решаемой задачи.

Фасет 1 (списочный) – Вид документа.

Значения фасета: акт, анкета, баланс, выписка, ведомость, договор, доверенность и т.д.

Фасет 2 (списочный) – Статус документа.

Значения фасета: действующий документ, документ, утративший силу, проект документа и т.д.

Фасет 3 (списочный) – Исполнитель документа.

Значения фасета: фамилии и инициалы должностных лиц и иных служащих Банка в алфавитном порядке.

Фасет 4 (текстовый) – Регистрационный номер и дата документа.

Значения фасета: валовые регистрационные номера и даты.

Емкость классификатора, построенного на основе фасетного метода классификации, зависит от числа фасетов и от количества признаков в фасетах.

При построении классификаторов на основе фасетного метода необходимо соблюдать следующие правила:

- значения признаков из различных фасетов не должны пересекаться;
- из всевозможных фасетов, характеризующих классифицируемое множество объектов, отбираются только существенные, то есть фасеты, обеспечивающие решение конкретных технико-экономических задач;
- фасеты должны занимать в классификаторе строго определенное место и иметь определенные идентификационные коды.

Фасетный метод построения классификационных группировок делает классификатор очень гибким, хорошо приспособленным для использования в условиях большой динамичности решаемых задач. При изменении характера задач или характеристик объектов классификации разрабатываются новые фасеты или дополняются новыми признаками уже существующие фасеты без коренной перестройки структуры всего классификатора.

В современных классификационных схемах часто одновременно используются оба метода классификации. Это обеспечивает возможность снизить влияние недостатков методов классификации и расширить возможность использования классификатора.