

Омский государственный технический университет

Кафедра «Электроснабжение промышленных предприятий»

Николаев М.Ю.

Информационные технологии в электроэнергетике

Системы управления базами данных.

Методы решения задач цифровой обработки
сигналов.

Методы решения задач анализа статистической
информации.

©ОмГТУ, 2015

Системы управления базами данных.

В течение многих лет преимущественно использовались плоские таблицы (плоские БД) типа списков в Excel. В настоящее время наибольшее распространение при разработке БД получили реляционные модели данных. Реляционная модель данных является совокупностью простейших двумерных таблиц – отношений (англ. relation), т.е. простейшая двумерная таблица определяется как отношение (множество однотипных записей объединенных одной темой). В реляционных БД используется несколько двумерных таблиц, в которых строки называются записями, а столбцы полями, между записями которых устанавливаются связи. Этот способ организации данных позволяет данные (записи) в одной таблице связывать с данными (записями) в других таблицах через уникальные идентификаторы (ключи) или ключевые поля.

Системы управления базами данных.

Принципы нормализации:

- В каждой таблице БД не должно быть повторяющихся полей;
- В каждой таблице должен быть уникальный идентификатор (первичный ключ);
- Каждому значению первичного ключа должна соответствовать достаточная информация о типе сущности или об объекте таблицы (например, информация об успеваемости, о группе или студентах);
- Изменение значений в полях таблицы не должно влиять на информацию в других полях (кроме изменений в полях ключа).

Системы управления базами данных.

Виды логической связи.

Связь устанавливается между двумя общими полями (столбцами) двух таблиц. Существуют связи с отношением «один-к-одному», «один-ко-многим» и «многие-ко-многим». Отношения, которые могут существовать между записями двух таблиц:

- один – к – одному, каждой записи из одной таблицы соответствует одна запись в другой таблице;
- один – ко – многим, каждой записи из одной таблицы соответствует несколько записей другой таблице;
- многие – к – одному, множеству записей из одной таблице соответствует одна запись в другой таблице;
- многие – ко – многим, множеству записей из одной таблицы соответствует несколько записей в другой таблице.

Системы управления базами данных.

Тип отношения в создаваемой связи зависит от способа определения связываемых полей:

- Отношение «один-ко-многим» создается в том случае, когда только одно из полей является полем первичного ключа или уникального индекса.
- Отношение «один-к-одному» создается в том случае, когда оба связываемых поля являются ключевыми или имеют уникальные индексы.
- Отношение «многие-ко-многим» фактически является двумя отношениями «один-ко-многим» с третьей таблицей, первичный ключ которой состоит из полей внешнего ключа двух других таблиц.

Системы управления базами данных.

Ключи. Ключ – это столбец (может быть несколько столбцов), добавляемый к таблице и позволяющий установить связь с записями в другой таблице. Существуют ключи двух типов: первичные и вторичные или внешние.

Первичный ключ – это одно или несколько полей (столбцов), комбинация значений которых однозначно определяет каждую запись в таблице. Первичный ключ не допускает значений Null и всегда должен иметь уникальный индекс. Первичный ключ используется для связывания таблицы с внешними ключами в других таблицах.

Внешний (вторичный) ключ - это одно или несколько полей (столбцов) в таблице, содержащих ссылку на поле или поля первичного ключа в другой таблице. Внешний ключ определяет способ объединения таблиц.

Из двух логически связанных таблиц одну называют таблицей первичного ключа или главной таблицей, а другую таблицей вторичного (внешнего) ключа или подчиненной таблицей. СУБД позволяют сопоставить родственные записи из обеих таблиц и совместно вывести их в форме, отчете или запросе.

Системы управления базами данных.

Существует три типа первичных ключей: ключевые поля счетчика (счетчик), простой ключ и составной ключ.

Поле счетчика (Тип данных «Счетчик»). Тип данных поля в базе данных, в котором для каждой добавляемой в таблицу записи в поле автоматически заносится уникальное числовое значение.

Простой ключ. Если поле содержит уникальные значения, такие как коды или инвентарные номера, то это поле можно определить как первичный ключ. В качестве ключа можно определить любое поле, содержащее данные, если это поле не содержит повторяющиеся значения или значения Null.

Составной ключ. В случаях, когда невозможно гарантировать уникальность значений каждого поля, существует возможность создать ключ, состоящий из нескольких полей. Чаще всего такая ситуация возникает для таблицы, используемой для связывания двух таблиц многие - ко многим.

Системы управления базами данных.

Программы, которые предназначены для структурирования информации, размещения ее в таблицах и манипулирования данными называются системами управления базами данных (СУБД). Другими словами СУБД предназначены как для создания и ведения базы данных, так и для доступа к данным. В настоящее время насчитывается более 50 типов СУБД для персональных компьютеров. К наиболее распространенным типам СУБД относятся: MS SQL Server, Oracle, Informix, Sybase, DB2, MS Access и т. д.

Методы решения задач цифровой обработки сигналов.

Цифровой обработкой сигналов принято называть в вычислительной технике арифметическую обработку последовательностей равноотстоящих вовремя отсчетов. Под цифровой обработкой понимают также обработку одномерных и многомерных массивов данных. По характеру изменения сигналы делятся на детерминированные и случайные. По своему смысловому определению детерминированные сигналы должны означать полностью известные сигналы на всем интервале их определения. Однако такие сигналы не несут новой информации и находят применение только в контроле и диагностике ЭВМ и систем обработки. Реальные сигналы всегда являются в какой-то мере неопределенными. Поэтому в дальнейшем при обработке сигналов под детерминированными сигналами понимаются такие, для которых известны законы их изменения во времени, но параметры, являясь также детерминированными, заранее могут быть неизвестными. Определение этих параметров и составляет существо ряда задач обработки сигналов.

Методы решения задач цифровой обработки сигналов.

На практике дискретные сигналы получают преимущественно из непрерывных. Процедура извлечения отсчетов дискретного сигнала из непрерывного называется дискретизацией по времени. Величину в этом случае называют интервалом (шагом) дискретизации. Её результатом является решетчатый сигнал. Равномерная дискретизация наиболее просто реализуется практически. Полученные отсчёты дискретного сигнала $x(i)$ для последующей обработки в ЭВМ квантуются по уровню, в результате чего получается цифровой сигнал. Поскольку квантование по уровню является нелинейной процедурой, то в общем случае результаты операций над дискретными сигналами не будут совпадать с результатами тех же операций над цифровыми сигналами. Расхождение результатов уменьшается с увеличением разрядности представления отсчетов цифрового сигнала.

Методы решения задач цифровой обработки сигналов.

Каждому дискретному сигналу могут соответствовать сколь угодно много непрерывных, совпадающих с ним в отсчетах, что говорит о многозначности задачи восстановления непрерывных сигналов по дискретным отсчетам. Такие непрерывные сигналы называются огибающими для данного дискретного. Простейшей огибающей является ступенчатый сигнал. Дискретные сигналы так же могут быть представлены в виде сумм либо четной и нечетной частей, либо постоянной и переменной составляющих. Цель обработки сигналов состоит в извлечении информации, а не энергии, тем не менее, энергия и мощность являются важнейшими характеристиками сигналов. Они позволяют, например, оценить степень взаимосвязанности различных сигналов и различных значений одного сигнала, проанализировать структуру сигнала, обусловленную внутренним распределением энергии или мощности, решить задачу выбора интервала дискретизации. Энергия и мощность могут служить мерой близости двух сигналов.

Методы решения задач анализа статистической информации.

В широком смысле статистика является наукой, изучающей массовые явления протекающие в совокупностях некоторых факторов или явлений определенного свойства и между взаимодействующими совокупностями. Сама же совокупность, как сумма фактов, признаков, явлений состоит из элементов, исчезновение одного из которых не уничтожает качественную характеристику этой совокупности. Так, население города остается его населением и после того, как одно из составляющих его содержания - физическое лицо переехало в другой город или другую местность или вообще покинуло данную страну. Или сельское хозяйство, транспорт и промышленность остаются определенными совокупностями соответствующими их характеристиками даже тогда, когда отраслевая структура или значимость их в производстве валового национального продукта претерпевает заметные изменения.

Методы решения задач анализа статистической информации.

Статистические методы — методы анализа статистических данных. Выделяют методы прикладной статистики, которые могут применяться во всех областях научных исследований и любых отраслях народного хозяйства, и другие статистические методы, применимость которых ограничена той или иной сферой. Имеются в виду такие методы, как статистический приемочный контроль, статистическое регулирование технологических процессов, надежность и испытания, планирование экспериментов. Дисперсионный анализ. Дисперсионный анализ применяется для исследования влияния одной или нескольких качественных переменных (факторов) на одну зависимую количественную переменную (отклик). В основе дисперсионного анализа лежит предположение о том, что одни переменные могут рассматриваться как причины (факторы, независимые переменные), а другие как следствия (зависимые переменные). Независимые переменные называют иногда регулируемыми факторами именно потому, что в эксперименте исследователь имеет возможность варьировать ими и анализировать получающийся результат.

Методы решения задач анализа статистической информации.

Основной целью дисперсионного анализа (ANOVA) является исследование значимости различия между средними с помощью сравнения (анализа) дисперсий. Разделение общей дисперсии на несколько источников, позволяет сравнить дисперсию, вызванную различием между группами, с дисперсией, вызванной внутригрупповой изменчивостью. При истинности нулевой гипотезы (о равенстве средних в нескольких группах наблюдений, выбранных из генеральной совокупности), оценка дисперсии, связанной с внутригрупповой изменчивостью, должна быть близкой к оценке межгрупповой дисперсии. Если вы просто сравниваете средние в двух выборках, дисперсионный анализ даст тот же результат, что и обычный t-критерий для независимых выборок (если сравниваются две независимые группы объектов или наблюдений) или t-критерий для зависимых выборок (если сравниваются две переменные на одном и том же множестве объектов или наблюдений).

Методы решения задач анализа статистической информации.

Сущность дисперсионного анализа заключается в расчленении общей дисперсии изучаемого признака на отдельные компоненты, обусловленные влиянием конкретных факторов, и проверке гипотез о значимости влияния этих факторов на исследуемый признак. Сравнивая компоненты дисперсии, друг с другом посредством F—критерия Фишера, можно определить, какая доля общей вариативности результативного признака обусловлена действием регулируемых факторов.

Исходным материалом для дисперсионного анализа служат данные исследования трех и более выборок, которые могут быть как равными, так и неравными по численности, как связными, так и несвязными. По количеству выявляемых регулируемых факторов дисперсионный анализ может быть однофакторным (при этом изучается влияние одного фактора на результаты эксперимента), двухфакторным (при изучении влияния двух факторов) и многофакторным (позволяет оценить не только влияние каждого из факторов в отдельности, но и их взаимодействие).

Методы решения задач анализа статистической информации.

Ковариационный анализ.

Ковариационный анализ — совокупность методов математической статистики, относящихся к анализу моделей зависимости среднего значения некоторой случайной величины одновременно от набора (основных) качественных факторов и (сопутствующих) количественных факторов. Факторы F задают сочетания условий, при которых были получены наблюдения X, Y , и описываются с помощью индикаторных переменных, причем среди сопутствующих и индикаторных переменных могут быть как случайные, так и неслучайные (контролируемые в эксперименте).

Если случайная величина Y является вектором, то говорят о многомерном ковариационном анализе.

Ковариационный анализ часто применяют перед дисперсионным анализом, чтобы проверить гомогенность (однородность, представительность) выборки наблюдений X, Y по всем сопутствующим факторам.

Контактная информация

Разработчик: к.т.н., доцент Николаев Михаил Юрьевич

Кафедра «Электроснабжение промышленных предприятий»

Адрес: пр. Мира, 11, корпус 6, кабинет 234

Тел.: 8(3812) 65-36-82

E-mail: MUNP@yandex.ru

Сайт кафедры: www.omgtu.ru