

АИС - СТАНДАРТ

Технологическая схема и программное обеспечение статистического анализа и моделирования показателей технического совершенства продукции для целей автоматизации экспертизы

Р 50-54-41-88

gost.e-ecolog.ru

ГОСУДАРСТВЕННЫЙ КОМИТЕТ СССР ПО СТАНДАРТАМ
(Госстандарт СССР)

Всесоюзный научно-исследовательский институт
по нормализации в машиностроении
(ВНИИНМАШ)

Утверждены

Приказом ВНИИНМАШ

№ 375 от I.I2.I987г.

АИС-стандарт

Технологическая схема и программное обеспечение статистического анализа и моделирования показателей технического совершенства продукции для целей автоматизации экспертизы

Р е к о м е н д а ц и и

Р 50-54-4I-88

Москва I988

УДК 681.3.6.37 Группа Т51

Р Е К О М Е Н Д А Ц И И

АИС-Стандарт
Технологическая схема и программное обеспечение статистического анализа и моделирования показателей технического совершенства продукции для целей автоматизации экспертизы

Р 50-54-41-88

ОКСТУ - 4099

В рекомендациях описаны технологическая схема и метод статистического анализа на этапах экспертизы для оценки технического совершенства продукции машиностроения. Приводятся структуры алгоритмического и программного обеспечения технологической схемы, предлагаются и описываются диалоговые программные процедуры этапов статистического анализа.

Излагаются принципы построения статистических банков данных.

Рекомендации рассчитаны на инженеров-экономистов, экспертов, математиков, программистов, занимающихся автоматизацией методов оценки технического совершенства продукции.

I. ТЕХНОЛОГИЧЕСКАЯ СХЕМА СТАТИСТИЧЕСКОГО АНАЛИЗА И МОДЕЛИРОВАНИЯ

I.I. Общие положения

Комплексное изучение технологических и технических систем в настоящее время выдвигает задачу создания автоматизированных программных систем по обработке экспериментальных данных, начиная с организации хранения собранной информации и кончая анализом результатов обработки данных. Разработка подобных программных комплексов требует разумного соединения проблемноориентированных банков данных (БнД) и специальных алгоритмов обработки, выполненных на основе конкретного математического аппарата, который в свою очередь определяется построенной моделью системы. Современный уровень развития программного обеспечения (ПО) предоставляет для этого широкие возможности. Разработаны СУБД, обладающие достаточной гибкостью и позволяющие строить разнообразные логические структуры данных, существуют пакеты прикладных программ, реализующие обширные математические алгоритмы, и т.п. Создание программных средств автоматизации различных задач стандартизации, в частности, экспертизы, позволяет повысить эффективность работ как самих экспертов, так и использования информации из создаваемых проблемноориентированных банков данных. Следует решительно переходить от пассивного накапливания информации к ее активному использованию путем аналитической обработки различными методами. Важнейшие методы обработки являются составной частью статистического анализа и моделирования.

I.2. Место статистического анализа и моделирования в формировании показателей технического совершенства продукции

Для определения места статистического анализа и моделирования в формировании показателей технического совершенства продукции и их оценки с целью создания автоматизированных систем экспертизы следует привести схематическое изображение цикла создания изделий машиностроения (в данном случае на примере электротехники) (рис. I.I). Из данной схемы вытекает, что влияние результатов статистического анализа комплексно и сохраняет значимость на различных этапах создания изделия: техническое задание, проектирование, анализ производственных возможностей, испытания, экспертиза и т.п. Локализовать место статистического анализа для формирования и анализа показателей технического совершенства продукции поможет схема рисунка I.2. Очевидно, что статистический анализ позволяет обеспечить решение задач в таких направлениях стандартизации, как экспертиза НТД, выделение групп однородной продукции, ранжирование изделий и показателей, прогнозирование взаимовлияния показателей, определение размеров выборок и интервалов испытаний и т.п. Поэтому создание программных средств реализации статистического анализа и моделирования, ориентированных на задачи стандартизации, в частности, экспертизы, является насущной необходимостью и требует *солидного методического обеспечения.*

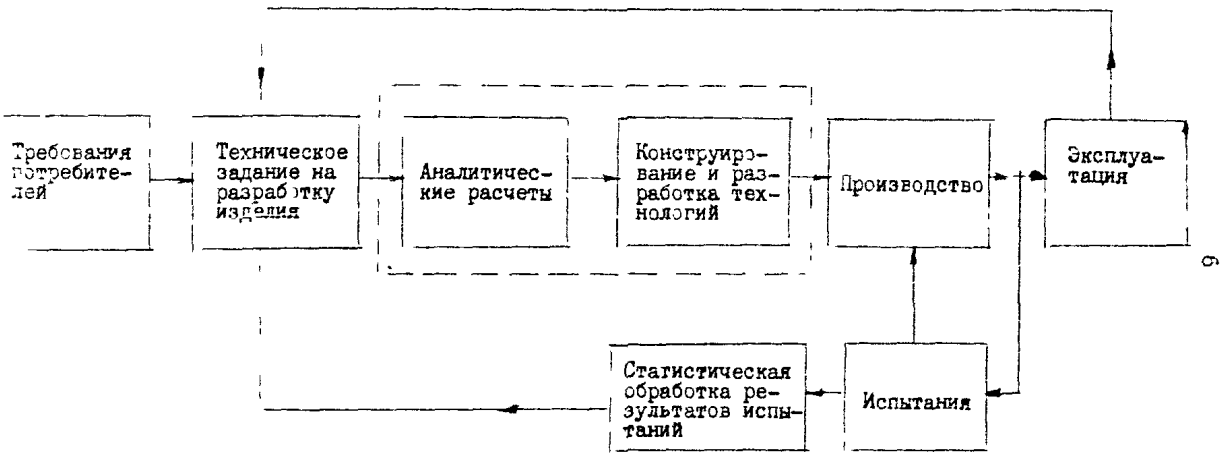


Рис. I.I. Схематическое изображение замкнутого цикла создания изделий электротехники

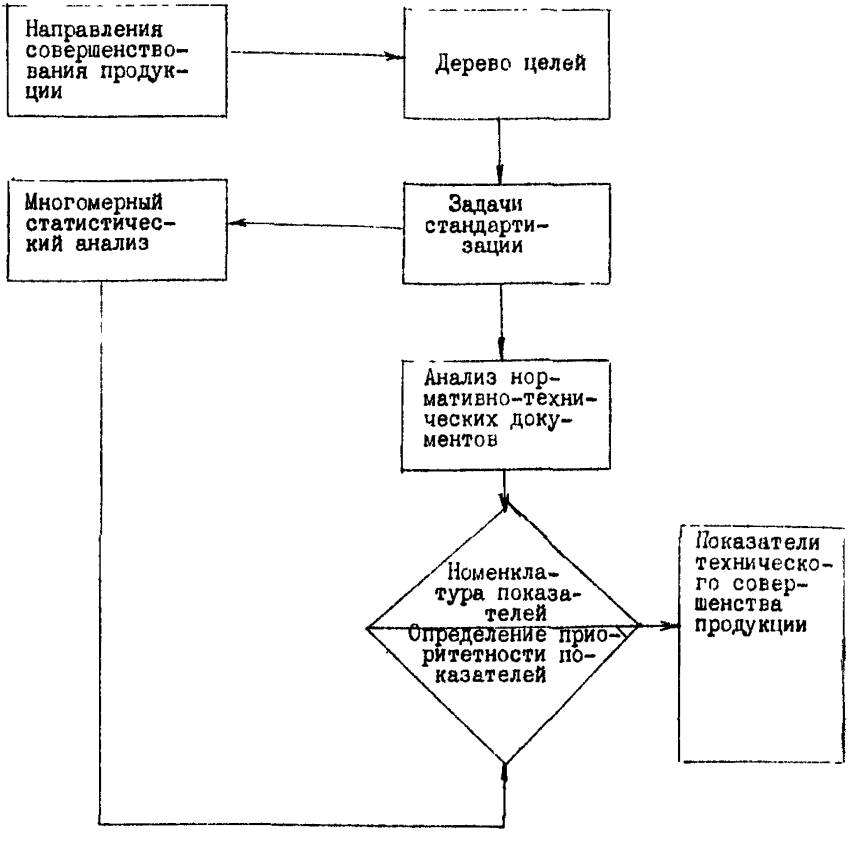


Рис. I.2. Схема формирования и анализа показателей технического совершенства электротехнической продукции.

I.3. Автоматизация моделирования как общая задача обработки данных

Автоматизация моделирования в общем виде предполагает переложение на ЭВМ следующих операций [2, 3] :

- ввод и формирование массива исходных данных для моделирования;
- приведение морфологического и функционального описания подсистем и элементов к требуемому для работы моделирующих алгоритмов виду;
- проведение имитационного эксперимента и управление им;
- обработка и анализ результатов имитации;
- управление генерацией имитационных программ, а также общее управление комплексом моделирования.

I.4. Технологическая схема статистического анализа и моделирования объектов стандартизации

Учитывая требования к автоматизации моделирования, стоящую перед данным исследованием задачу и общую схему формирования показателей технического совершенства, можно предложить следующую технологическую схему моделирования (рис. I.3). В свою очередь данная схема отражает первый этап более сложного и трудоемкого процесса проектирования и реализации имитационной системы моделирования объектов стандартизации. Создание таких многоэтапных программных комплексов приводит к необходимости обеспечивать преемственность как многих процедур подпрограмм, так и рабочих файлов и баз данных (БД). Принципы создания БД, их

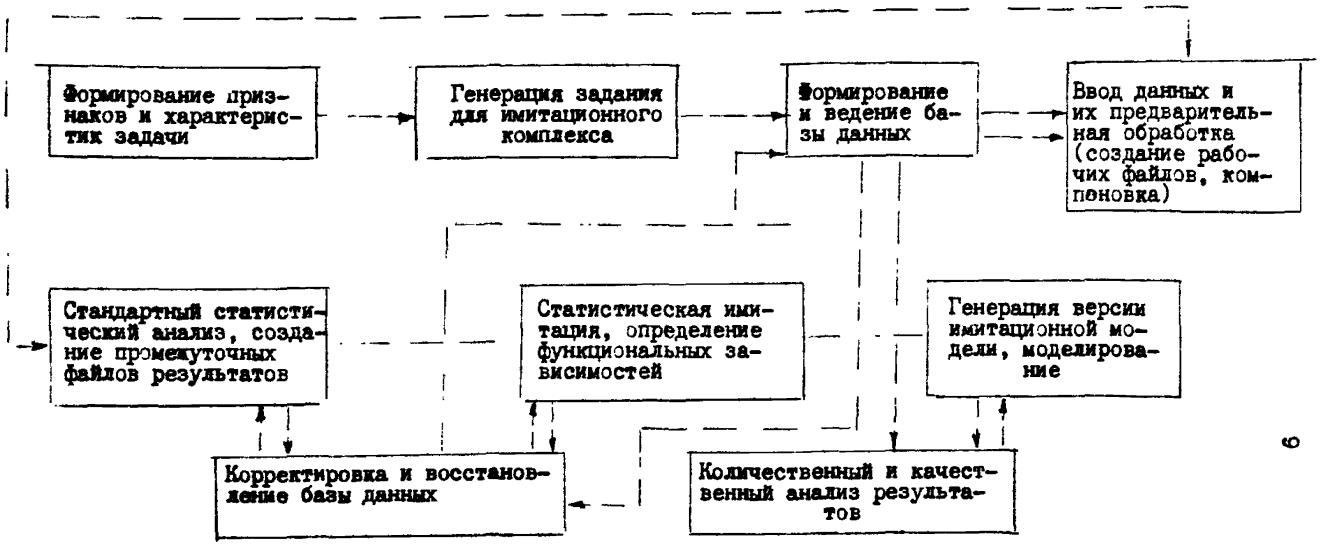


Рис. I.3. Технологическая схема статистического анализа и моделирования объектов стандартизации.

ведения хорошо изучены и разработаны [6] . Основные требования к базам данных сводятся к следующему:

- обеспечение возможности быстрого поиска данных в БД и представление их в форме, удобной для дальнейшей обработки;
- оперативный контроль данных;
- простота корректировки БД без изменения всего комплекса прикладных программ;
- простота и гибкость доступа к данным, создание формально независимых алгоритмов доступа к данным;
- сокращение избыточных затрат памяти на физическом носителе, а также времени на корректировку БД;
- минимизация времени доступа к данным (времени ответа пользователю).

I.5. Общая структура программного обеспечения

Как видно из рисунка I.4, все программное обеспечение подразделяется на четыре части: формирования и ведения БД, статистического анализа, статистической имитации (моделирование), анализа результатов. Программы формирования и ведения БД, их состав определяются логической и физической структурой БД, которые в свою очередь зависят от конкретного представления описания объекта стандартизации, используемых технических средств, математического обеспечения ЭВМ. Программы статистического анализа представлены в основном программами пакета стандартного статистического анализа и специальной статистики, а также процедурами формирования рабочих файлов, необходимых при расчете. Две другие части ПО представлены моделирующими программами и

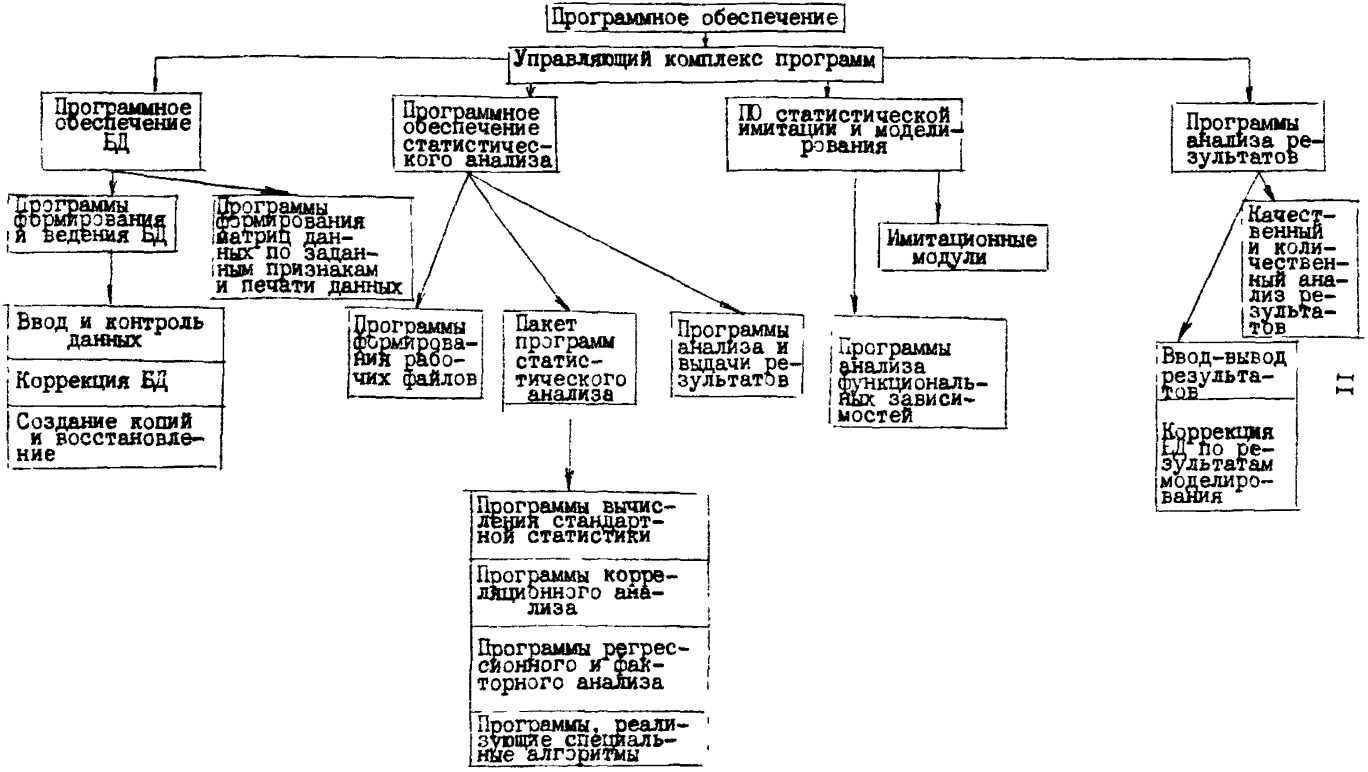


Рис. I.4. Общая структура программного обеспечения статистического анализа и моделирования объектов стандартизации.

средствами сервиса и анализа результатов. Несмотря на существование различных БД и СУБД, для частных задач стандартизации целесообразно разрабатывать оригинальные БД, более простые и ориентированные на конкретные задачи, так как в большинстве случаев БД не является самоцелью, объем хранимой информации не столь велик, а сложность программного обеспечения даже среднего банка данных [6, 7] весьма значительна. Проблемноориентированные БД должны обладать одним важным свойством - стандартным интерфейсом, что обеспечит возможность их стыковки или включения в большие банки данных. Программы контроля данных обязаны учитывать тот факт, что алгоритм контроля обычно формализуется и закладывается в качестве справочной информации. В простейшем варианте это может быть проверка величин на принадлежность некоторым диапазонам.

I.6. Понятие статистической имитации как вида моделирования

Следует определить понятие статистической имитации. Под статистической имитацией будем в дальнейшем понимать имитационное моделирование, которое отвечает ряду основных требований и положений теории имитации [2, 3] , но при котором вместо функциональных зависимостей используются количественные статистические оценки процессов в объекте моделирования на уровне взаимодействия подсистем и элементов в структурной схеме объекта моделирования.

Отметим ряд важных черт статистической имитации, которые

необходимо учитывать. Во-первых, любой статистический метод является достаточно грубым при оценке причинно-следственных связей в объекте. Однако здесь надо помнить, что такой предварительный этап моделирования необходим, особенно при слабой изученности объектов исследования, функциональных связей их компонентов. Во-вторых, известно, что любые статистические методы дают при прогнозе расширяющийся доверительный интервал, что, естественно, накладывает ряд ограничений на область использования результатов такого анализа. При статистической имитации имеется возможность учесть данную особенность перебором комбинаций интервалов варьирования основных факторов. В-третьих, статистическая имитация позволяет моделировать качественные изменения в системе и подойти к формулировке функциональной модели. Имитационные программы используют усредненные статистические оценки, а не точные зависимости. Это позволяет чаще всего выделять по результатам моделирования не численные значения, а области значений факторов ("пороги"), при которых можно ожидать качественные изменения в функционировании системы. В-четвертых, разработка специальных программ статистической имитации, базирующихся на стандартных статистических методах (см. напр. [I, 4]), позволяет при сохранении принятых в работе принципов автоматизации имитационного моделирования подойти к реализации элементов автоматизации как ведения БД, так и самого моделирования. Автоматизация моделирования требует стандартной математической схемы, лежащей в основе описания элемента структурированной схемы. Разрабатываемый метод статистической имитации учитывает это требование на уровне проведения имитационного эксперимента и генерации имитационной программы.

Процесс статистической имитации воссоздает на ЭВМ общесистемную среднестатистическую ситуацию, и, как любой подход к имитационному моделированию, предполагает проведение серий машинных экспериментов для разных внешних и внутренних условий функционирования объекта моделирования. В первом приближении можно сказать (см. рис I.4) что все программное обеспечение делится на управляющие собственно имитационные и сервисные группы программ. Это важно при описании структуры алгоритмического обеспечения комплекса программ статистической имитации.

I.7. Структура алгоритмического обеспечения статистического анализа и моделирования

Алгоритмы управляющих модулей (блок I рис. I.5) реализуют генерацию модели данной конфигурации на основе морфологического описания системы (унифицированные структурные схемы) 2-4 ,планируют использование языка управления заданиями операционной системы, определяют формы представления различных данных для работы имитационных программ и т.п. Алгоритмы 2, 3 и 4 блоков схемы рис.I.5 в основном реализуют принципиальную технологию формирования и ведения баз данных [6, 7] . Алгоритмы 5 блока подробно изложены в литературе [3,4,6,7]. Алгоритмы блоков I, 6, 7 в достаточной степени оригинальны, поэтому иллюстрации на примерах этих типов алгоритмов будет посвящено дальнейшее изложение (алгоритмы блока 8 могут быть детализованы пока в основном лишь для конкретной задачи и объекта, однако в общей схеме их необходимо учитывать).

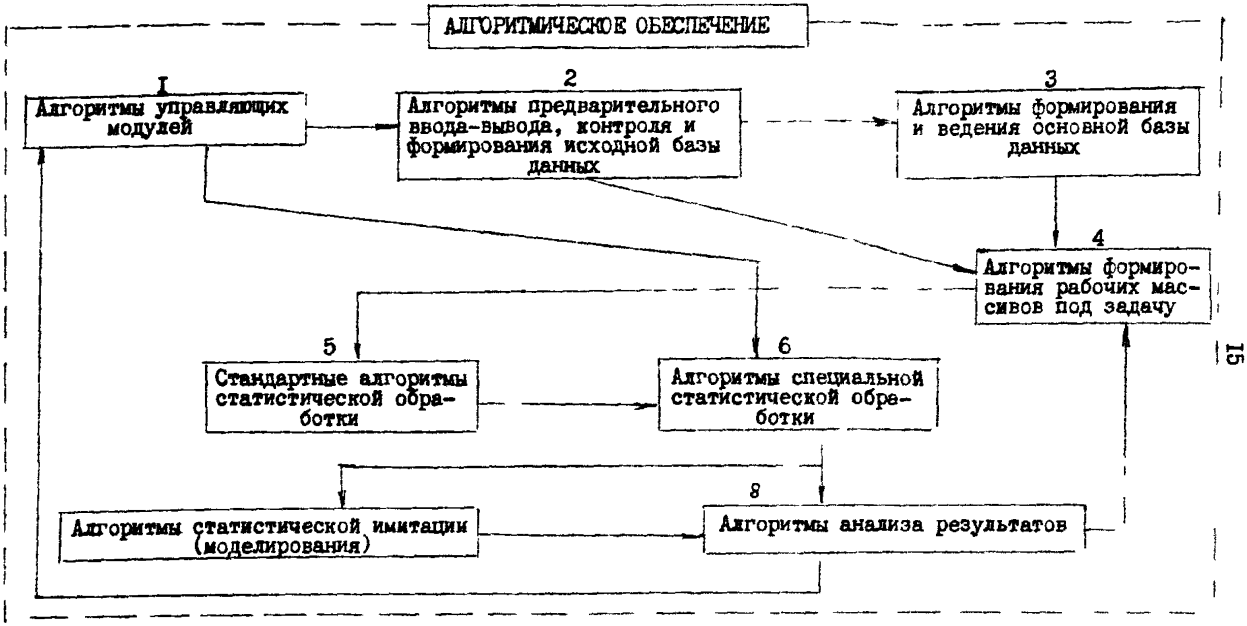


Рис. I.5. Алгоритмическое обеспечение статистического анализа и моделирования и возможные связи его блоков

Приводимые примеры алгоритмов должны носить достаточную степень общности, так как изложение подробных блок-схем на данном уровне описания алгоритмического обеспечения привело бы лишь к непомерному увеличению объема без качественного выигрыша в информативности.

I.8. Алгоритм управляющего модуля

Шаг I. Настройка программного комплекса (задается по бланку описания), в которой указывается число и тип подсистем и элементов в соответствии с морфологическим описанием. Также задаются характеристики задачи исследования объекта (например, время моделирования, степень требуемой точности прогноза и т.п.).

Шаг 2. Автоматическое формирование языком управления заданиями пакета требований на определение и распределение различных типов памяти операционной системы для данного модельного исследования. Формирование такого пакета может осуществиться с помощью базы данных символьных констант.

Шаг 3. Внесение конкретных данных в рабочие файлы.

Шаг 4. Формирование по исходному описанию структуры имитационной программы, получение и проверка таблицы наличия сгенерированных имен необходимых модулей.

Шаг 5. Проверка наличия фактических параметров и файлов данных в базе данных для списка имен сгенерированных модулей. Формирование запроса на дополнение базы данных и разрешение ссылок на имена требуемых модулей и наборов данных.

Шаг 6. Обмен с базой данных и заполнение общих областей памяти. Контроль и печать исходного и объектного модулей созданной имитационной программы.

Шаг 7. Передача управления модулю статистической имитации. Проведение имитационного эксперимента.

Шаг 8. Работа модуля анализа результатов.

Шаг 9. Общая печать и обновление базы данных результатами имитации. Задание в диалоговом режиме управляющему модулю условий продолжения работы (тогда переход к шагу 2) или прерывания задания.

Шаг 10. Закрытие наборов данных, внесение информации в справочники ведения базы данных. Завершение работы.

I.9. Алгоритм одного из модулей специальной статистики (расчет комплексного влияния факторов)

Шаг I. Ввод исходной матрицы и представление в виде, требуемом для расчета.

Шаг 2. Расчет матрицы коэффициентов парной и частной корреляции, а также коэффициентов достоверности Стьюдента.

Шаг 3. Расчет коэффициентов множественной корреляции, детерминации, регрессии, дисперсий и другой стандартной статистики.

Шаг 4. Получение значений "порогов" с помощью специальной процедуры для каждого исследуемого фактора.

Шаг 5. Получение редуцированной матрицы из входной матрицы наблюдений для расчета по порогам значений коэффициентов удельного влияния данного фактора на дисперсию функции отклика.

Шаг 6. Пересчет необходимых вспомогательных величин для редуцированной матрицы.

Шаг 7. Расчет коэффициентов удельного влияния факторов для "порогов" данного фактора.

Шаг 8. Если расчет произведен для всех факторов, то печать и переход к шагу 9, иначе к шагу 5.

Шаг 9. Пополнение базы данных, передача управления головному модулю.

I.IO. Алгоритм статистической имитации

Как любая модель есть упрощенное представление объекта исследования, так и любой моделирующий алгоритм есть упрощенное представление всей сложной совокупности процессов функционирования. Степень упрощения и ее допустимые пределы определяются задачей, которую призван решать какой-либо метод моделирования. Помимо статистических ограничений (непосредственно статистического анализа) в данном случае представляется целесообразным назвать ряд таких упрощений, содержащихся в предлагаемом алгоритме статистической имитации. Во-первых, предполагается, что реакция модельной системы на сходные возмущения в случае, если состояние системы и предыстория ее развития сходны с наблюдениями в натуре, также одинакова. Во-вторых, количественные изменения функции отклика (или ее дисперсии) пропорциональны весам коэффициентов удельного влияния факторов (естественно, в интервалах значений факторов, имеющихся в натурных измерениях). В-третьих, очевидно допустить при сравнении двух состояний объекта, что количественные расхождения, соизмеримые с ошибкой натурного наблюдения или измерения, незначимы.

Алгоритмы представляют собой логическую последовательность действий, при точном выполнении которой можно решить поставленную задачу изучения объекта. Для имитационного моделирования, в особенности в "игровых" задачах, характерно использование в основе алгоритмов эвристических процедур. Такие алгоритмы носят название логико-эвристических алгоритмов. Приводимый ниже алгоритм относится к этому типу.

Шаг I. Формирование вектора состояния имитируемого компонента системы на основе морфологического описания по матрице состояния (У) системы. Занесение во вспомогательный файл требуемых для работы строк матрицы конечных приращений.

Шаг 2. Задание возмущения в системе и расчет вектора входных сигналов в имитируемый компонент (число координат равно числу воздействующих факторов на данный модельный момент).

Шаг 3. Розыгрыш фиксированных значений отклонений по X (факторы) с помощью задаваемых плотностей распределения. Расчет матрицы относительных отклонений по X и создание вспомогательного файла для анализа сходства модельного состояния компонента и данных натурного эксперимента.

Шаг 4. Выбор всех комбинаций отклонений по X, для которых натурный и модельный результаты можно принять сходными.

Шаг 5. Реализация процедуры анализа идентичности вектора модельного и натурного состояний и реакций системы. Если обнаружены идентичные вектора, то переход на шаг I0, иначе на шаг 6.

Шаг 6. Анализ расхождений модельного вектора состояний и матрицы, созданной по данным натурного наблюдения. Формирование массива индексов для координат, по которым расхождения значительны.

Шаг 7. По массиву индексов расхождений организация цикла для работы процедуры коррекции. Эта процедура по функции для коэффициента удельного влияния данного фактора в зависимости от значений фактора рассчитывает численное значение величины коррекции дисперсии У. Такой расчет производится с учетом имеющихся комбинаций отклонений воздействующих факторов в матрице натурных наблюдений и численных значений коэффициентов удельного влияния данного фактора в определенных специальной процедурой областях значений этого коэффициента по реализациям для вариации различных факторов.

Шаг 8. Если отличие вектора состояния моделируемой системы значимо по нескольким факторам одновременно, то определяются области коэффициентов удельного влияния для соответствующей комбинации отклонений факторов. Затем реализуется процедура розыгрыша фиксированных значений коэффициентов удельного влияния (средних для полученных областей значений), и вновь анализируется вспомогательная матрица на сходство комбинаций вектора состояний. Далее в случае сходства процедура аналогична шагу 7, иначе шаг 8 повторяется снова.

Шаг 9. Полученные таким образом значения строк для У и Х добавляются к значениям исходной матрицы и производится тщательный анализ новой матрицы на достоверность

прогноза. Если достоверность подтверждается, то переход к шагу I, иначе к шагу 3.

Шаг IO. После окончания имитации по созданной матрице рассчитываются области абсолютного влияния какого-либо фактора и анализируется соответствие имеющимся функциональным зависимостям (адекватность проведенной статистической имитации). Если адекватность количественная и качественная подтверждается, то определяются возможные новые функциональные зависимости по абсолютным областям влияния отдельных факторов на функции отклика.

Шаг II. Пополнение базы данных результатами имитации. Уничтожение вспомогательных наборов. Передача управления вызывающему модулю. Печать.

Итак, представленное алгоритмическое обеспечение программного комплекса статистической имитации охватывает все этапы модельного исследования объекта стандартизации от ввода и преобразования исходных данных, формализуемых специальными бланками, до качественного анализа результатов имитационного эксперимента. Изложенные выше типовые алгоритмы являются хорошей иллюстрацией многоэталной статистической имитации.

2. АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ ОЦЕНКИ ТЕХНИЧЕСКОГО СОВЕРШЕНСТВА

2.I. Постановка задачи создания конкретных автоматизированных программных систем

Задача автоматизации экспертизы должна решаться поэтапно,

причем на каждом этапе эксперт должен получать бо́льшие возможности для проведения качественной оценки технического совершенства НТД и передавать ЭВМ часть рутинных операций процесса экспертизы. Следует выделить следующие основные этапы автоматизации задач экспертизы.

I-ый этап. Разработка и внедрение банков данных о техническом совершенстве продукции. На этом этапе выявляются требования пользователя-эксперта к информационному обслуживанию процесса оценки технического совершенства НТД, определяется перечень входных документов (источников информации), устанавливается структура информации и перечень реквизитов. При внедрении таких автоматизированных информационных систем пользователь-эксперт получает возможность обобщать информацию по своей предметной области, средствами системы управления базами данных (язык запросов и др.) получать сведения по сформулированным запросам, проводить простую арифметическую и логическую обработку информации, содержащейся в базе данных. В системах типа "банк данных" информация независима от программных средств, и это создает предпосылки для перехода к последующим этапам автоматизации задач экспертизы.

2-ой этап. Разработка и формирование базы методик обработки данных, которая является надстройкой над базой данных о техническом совершенстве продукции. На этом этапе выявляются основные задачи экспертизы, определяются возможные методы решения этих задач, разрабатываются программные средства, дающие возможность эксперту пользоваться методиками обработки информации.

3-ий этап. Создание экспертных систем оценки технического совершенства НТД продукции. Этот этап не связан непосредственно с двумя предыдущими и предполагает разработку программного комп-

лекса, моделирующего поведение эксперта при решении задач оценки в конкретной предметной области, характеризуемого такими особенностями: I) решения обладают ясностью, высоким качеством и требуют минимума ресурсов; 2) решения являются результатом применения символических рассуждений, базирующихся на эвристиках; 3) система способна анализировать и объяснять свои действия и знания; 4) система способна приобретать от пользователя (эксперта) новые знания и менять в соответствии с ними свое поведение. Экспертная система состоит из следующих компонентов: решателя, базы данных, базы знаний, приобретения знаний, объяснения и диалогового.

В настоящее время очевидно, что при возрастающих объемах перерабатываемой информации и при повышении требований к качеству решений экспертов на первый план выходят задачи разработки программных систем оценки технического совершенства НТД продукции.

2.2. Существующие схемы оценки технического совершенства

В настоящее время применяются следующие основные подходы к оценке технического совершенства.

В основу оценки технического совершенства НТД продукции кладется сравнение совокупности показателей качества этой продукции с соответствующей совокупностью показателей качества базового образца (совокупность базовых значений показателей).

При оценке технического совершенства применяют дифференциальный, комплексный и смешанный методы.

Дифференциальный метод основан на использовании единичных

показателей качества продукции. При этом определяют, достигнут ли уровень базового образца в целом, по каким показателям он достигнут, какие показатели наиболее сильно отличаются от базовых. При дифференциальном методе рассчитывают относительные показатели качества продукции по формулам:

$$q_i = \frac{P_i}{P_{i\delta}} \qquad q_i^1 = \frac{P_{i\delta}}{P_i} \qquad (i = 1, \ldots n)$$

где P_i – значение i-го показателя качества оцениваемой продукции; $P_{i\delta}$ – значение i-го базового показателя; n – количество показателей качества продукции.

При расчете выбирают ту формулу, при которой увеличению относительного показателя отвечает улучшение качества продукции.

В случае, когда часть значений относительных показателей больше или равна единице, а часть – меньше единицы, применяют комплексный или смешанный метод оценки технического совершенства НТД продукции.

Комплексный метод оценки технического совершенства основан на применении обобщенного показателя качества продукции.

Обобщенный показатель представляет собой функцию от единичных показателей качества. Он может быть выражен:

– главным показателем, отражающим основное назначение продукции;

– интегральным показателем качества продукции;

– средним взвешенным показателем.

Когда имеется необходимая информация, определяют главный показатель и устанавливают функциональную зависимость его от исходных показателей.

Смешанный метод оценки технического совершенства НТД продукции основан на совместном применении единичных и комплексных

(групповых) показателей. Смешанный метод применяют в случаях:

- когда совокупность единичных показателей качества достаточно обширна и дифференциальный метод не дает обобщающих выводов;

- когда комплексный показатель качества в комплексном методе не отражает все существенные свойства продукции.

При смешанном методе оценки выполняют следующие действия:

- часть единичных показателей объединяют в группы и для каждой группы определяют соответствующий комплексный (групповой) показатель. Отдельные важные показатели не объединяются в группы и используются в дальнейшем анализе как единичные;

- на основе полученной совокупности комплексных и единичных показателей оценивают уровень технического совершенства НТД продукции дифференциальным методом.

2.3. Технология создания диалоговой системы автоматизированного проектирования методик оценки[x)]

Различные стратегии решения задач оценки и общая схема решения подробно рассмотрены в работе [5]. В настоящих рекомендациях для решения задачи повышения эффективности работ по оценке технического совершенства продукции предлагается диалоговая система автоматизированного проектирования методик оценки (САПРМО), составными частями которой являются методы обработки информации. Данная система позволяет учитывать личный опыт эксперта и основана на информационной базе данных о техническом совершенстве продукции.

Прежде всего при разработке САПРМО необходимо выделить основные уровни такой системы (рис. 2.1 [5]).

На первом уровне находятся основные задачи обработки инфор-

x) Разделы 2.3 и 2.4 подготовлены при участии к.т.н. М.Ш.Лернера

мации при экспертизе. Эти задачи экспертизы определяются при помощи экспертов, на основе анализа различных руководящих документов и другими способами.

На втором уровне - задачи формализованной поддержки, позволяющие решать основные задачи экспертизы НТД.

На третьем уровне - конкретные методы (формальные модели решения), реализующие задачи 2-го уровня.

Матрица взаимосвязи задач I-го и 2-го уровня представлена на рис. 2.3, матрица взаимосвязи задач 2-го уровня и формальных моделей решения - на рис. 2.2.

Общая схема диалоговой программы приведена на рис. 2.4. Эксперт формирует методику оценки ТУ в ходе диалога с ЭВМ, работая непосредственно за дисплеем. При разработке диалоговой программы необходимо учитывать различный уровень подготовки экспертов и соответственно предусматривать возможность вывода на экран различных справок, сведений, инструкций, облегчающих общение эксперта с машиной.

Еще до формирования методики оценки эксперт вводит с экрана дисплея информацию об аттестуемой продукции в базу данных.

В частном случае эксперт может проводить анализ информации, уже имеющейся в базе данных. Затем эксперт определяет (из меню) типовую задачу экспертизы. Отвечая на вопросы диалоговой программы, вводя параметры или выбирая одну из предложенных альтернатив, эксперт определяет выбор задач поддержки. Выбор формальных моделей решения может осуществляться также при помощи эксперта или автоматически. После того, как ЭВМ проанализирует информацию по сформированной в ходе диалога методике, эксперт оценивает результаты анализа, и, если результат его удовлетворяет, методика

I	Основные задачи обработки информации при экспертизе НТД	- поиск аналогов; - определение номенклатуры показателей, обладающей максимальной информационностью; - выбор базовых образцов; - определение лучшего отечественного (зарубежного) аналога; - определение места оцениваемого изделия в ряду аналогов
П	Задачи формализованной поддержки	- группирование; - ранжирование; - статистическая обработка; - ...
Ш	Формальные модели решения	- методы кластерного анализа; - статистические методы (корреляционные, дисперсионные, факторные); - методы ранжирования

Рис. 2.I. Основные уровни автоматизированной системы оценки НТД

Задачи формализованной поддержки / Формальные модели решения	Кластерный анализ	Факторный анализ	Корреляционный анализ	Метод Парето	Метод Электра
ГРУППИРОВАНИЕ	.		.		
РАНЖИРОВАНИЕ		.		.	.
СТАТИСТИЧЕСКАЯ ОБРАБОТКА	.	.	.		
ОПРЕДЕЛЕНИЕ ПРИОРИТЕТНОСТИ ПОКАЗАТЕЛЕЙ	.	.	.		
ПОДБОР АНАЛОГОВ	.		.		

Рис.2.2. Матрица (2) связи задач формализованной поддержки и формальных моделей решения

Основные задачи обработки информации при экспертизе НТД \ Задачи формализованной поддержки	Группирование	Ранжирование	Статистическая обработка
Поиск аналогов	•		
Определение приоритетности показателей, обладающей максимальной информационностью	•		•
Выбор базовых образцов	•		
Определение лучшего отечественного (зарубежного) аналога		•	
Определение места оцениваемого изделия в ряду аналогов	•	•	
Оценивание интервалов значимости показателей технического совершенства			

Рис. 2.3. Матрица (I) связи основных задач обработки информации и задач формализованной поддержки

фиксируется в банке методик оценки для данного эксперта и данного вида продукции. В противном случае эксперт формирует другую методику оценки. Подобный подход выработки стратегии решения из различных сочетаний алгоритмов и процедур преобразования данных показан в работе на примере решения задачи стратификации [5].

Структура диалоговой программы формирования методики оценки приведена на рис. 2.5. Программа – монитор реализуется на языке программирования высокого уровня (ПЛ-I, язык описания сценария диалога СУБД ИНЕС).

В задачи программы-монитора входит ввод и вывод информации на экран дисплея, вызов блоков обработки информации и в конце запись в банк методик оценки совокупности ключей и параметров, определяющих состав и последовательность блоков обработки.

Блоки I, N – загрузочные модули, реализующие различные алгоритмы типовых формальных моделей решения. Разные блоки могут реализовываться на разных языках программирования: АССЕМБЛЕРЕ, ПЛ-I, ФОРТРАНЕ. Такой подход придает системе гибкость, позволяет включать в САПРМО разработанные независимо программы ранжирования, группировки и др.

2.4. Формальная модель автоматизированной оценки

Для описания процесса проектирования в САПРМО методики оценки НТД используются типовые структурные модели. Элементами таких моделей являются проектные процедуры и операции, представленные в виде модулей программного обеспечения. Каждый из этих модулей реализует различные формальные модели решения (рис. 2.I и 2.5).

Типовая структурная модель включает матрицу связи элементов и их характеристик $F[M] = [F \times M]$ (F – характеристики моду-

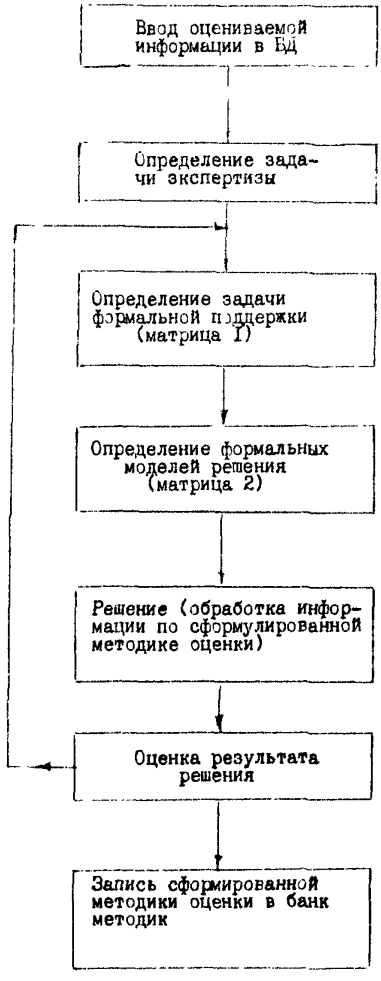


Рис. 2.4. Общая схема диалоговой программы анализа информации НТД

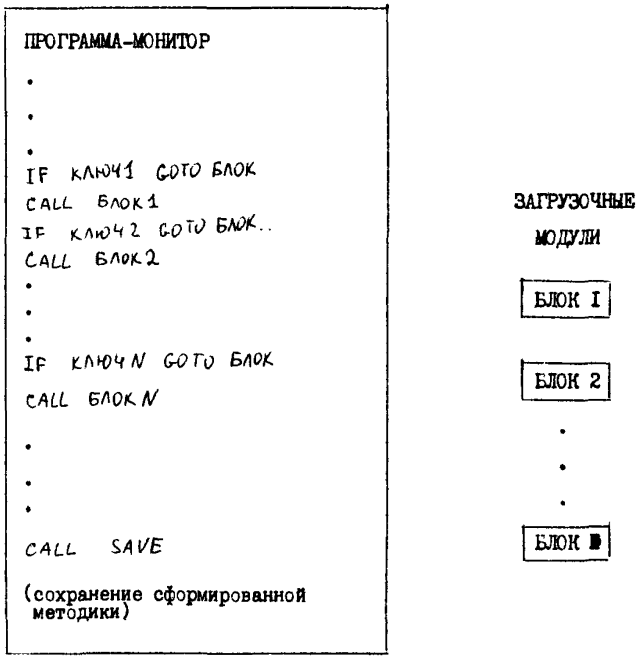


Рис. 2.5. Структура диалоговой программы формирования методики оценки

лей, M - модули), граф связи модулей системы программного обеспечения $G=(M,C)$, а также количественную модель, позволяющую выбрать оптимальный вариант методики оценки [8].

Структурные связи между элементами графа $G=(M,C)$ определяются на основе анализа возможных последовательностей выполнения проектных процедур и операций. В качестве количественной модели для выбора оптимальной методики оценки может служить, например, таблица модулей программного обеспечения и процессорного времени на их выполнение.

Эксперт, работая с диалоговой программой за дисплеем, определяет характеристики необходимой ему методики оценки. Исходя из этих характеристик по матрице $F(M)$ определяется состав всех модулей, реализующих указанные экспертом требования. Затем по графу $G=(M,C)$ строятся различные варианты методик, отличные по составу и последовательности выделенных на предыдущем этапе модулей программного обеспечения. По количественной модели определяется оптимальная (например, с точки зрения быстродействия) методика оценки.

3. ПРИНЦИПЫ ПОСТРОЕНИЯ СТАТИСТИЧЕСКИХ БАНКОВ ДАННЫХ И СТРУКТУРЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

3.1. Постановка задачи создания статистических банков данных

Для управления техническим совершенством продукции и его оценки экспертизы НТД требуется информация, содержащаяся в разных документах (функционирующих в системе Госстандарта и в отраслях), таких как: отраслевые карты технического уровня, протоколы испытаний, государственные и отраслевые стандарты,

технические условия, промышленные каталоги, международные и национальные стандарты. Все эти документы разобщены, и поэтому доступ к данным, содержащимся в них, представляет значительные трудности, а оперативный доступ вообще невозможен.

Указанные трудности позволяет преодолеть создание автоматизированных информационных систем (АИС), содержащих наиболее важные сведения:

- библиографические данные о документе, из которого взята информация (обозначение и наименование документа, дата регистрации и т.д.);

- библиографические данные о продукции (предприятия или фирма, страна, даты начала и конца изготовления и т.д.);

- фактографические данные о продукции (значения показателей качества и технического уровня - важнейшая часть всех данных).

Концентрация данных из документов в одной или нескольких связанных между собой АИС позволяет организовать оперативный доступ к данным для управления техническим уровнем продукции. Реализация информационного обеспечения работ экспертизы НТД состоит из двух направлений работ:

- информационного обслуживания экспертов данными из информационной системы;

- использование данных информационной системы для работы пакетов прикладных программ, осуществляющих их статистическую обработку и выдачу рекомендаций экспертам.

3.2. Виды программных средств

Для создания АИС возможно использование программных средств

трех видов:

- языков программирования;
- систем управления базами данных (СУБД);
- языков моделирования (ЯМ).

Первый вид предполагает разработку программ на языках программирования (ПЛ -I, ФОРТРАН, АССЕМБЛЕР и др.), осуществляющих описание, ввод и манипулирование данными. Недостаток такого подхода заключается в зависимости программ от данных, что вызывает необходимость их изменения при реорганизации АИС, удаление некоторых реквизитов, добавление новых, изменение описаний реквизитов .

Второй вид программных средств включает разработку программ на языках, входящих в СУБД (описания структур данных, входных и выходных машинно-ориентированных форм, манипулирования данными), в которых обращение к данному уже осуществляется не по адресу, а по имени, присвоенному в описании данных. Поэтому при изменении некоторых данных в процессе функционирования системы программы не меняются. Данный подход является гораздо более гибким по сравнению с первым. Помимо гибкости и независимости программ от данных второй вид программных средств позволяет значительно сократить сроки разработки АИС за счет того, что языки, входящие в состав СУБД, ориентированы на описание, обработку и доступ к данным, что делает их проблемно-ориентированными языками очень высокого уровня. Программирование на них - процесс гораздо менее трудоемкий, чем программирование на универсальных языках (ПЛ -I, ФОРТРАН, АССЕМБЛЕР и др.).

3.3. Программные средства расширения возможностей СУБД

Создание на основе СУБД автоматизированных банков данных, являющихся ядром АИС и содержащих фактографические данные о продукции, позволяет решить проблему информационного обслуживания пользователей (экспертов). Однако автоматизированная обработка данных пакетами прикладных программ статистической обработки представляет некоторые трудности, состоящие в том, что формат входных данных индивидуален для каждого пакета. Поэтому для проведения статистической обработки необходимо осуществить выборку информации из БД и организовать ее ввод в нужном формате в ППП вручную. При условии многоразового решения задач, а также требования автоматизированной обработки данных подобное положение неприемлемо.

Решение указанных проблем связано с созданием программных средств, расширяющих возможности СУБД, которые позволяют реализовать:

- форматирование выбранных данных из базы для задач аналитической обработки с использованием программ-адаптеров;
- введение численных результатов выборки из базы в промежуточный массив хранения;
- организацию сопряжения между программами выборки данных из базы и программами статистической обработки (выход программ выборки после обработки программами-адаптерами направляется автоматически на вход программ обработки);
- выдачу машинограмм и графиков, характеризующих сводные результаты аналитической обработки, на АЦПУ и графопостроитель;
- хранение численных результатов работы аналитических программ и программ моделирования.

Любой СУБД предоставляет возможности:

- доступа к данным;

- обработки данных.

Возможности доступа и обработки различны у разных СУБД, но обычно обработка данных в них ограничивается информационно-справочными задачами. Развитая статистическая обработка данных (агрегирование, моделирование) предполагает наличие статистической или обычной СУБД, дополненной возможностями статистической обработки данных, в которых наряду с подсистемами описания, ввода и вывода данных (информационные задачи) присутствует подсистема статистической обработки данных.

Подсистема статистической обработки должна включать:

- все возможности обработки данных, предоставляемые статистическими пакетами;

- удобные средства доступа к возможностям статистической обработки данных (в виде диалоговых языков).

В статистических пакетах основной упор делался на процедуры обработки данных, доступ к которым осуществлялся в пакетном режиме, при этом входной набор данных имел жесткий формат. В статистических СУБД (базах данных) главное внимание уделяется "дружественному" интерфейсу пользователя с процедурами статистической обработки при сохранении всех возможностей статистических пакетов.

Процедуры статистической обработки включаются в БнД различными способами:

- включение процедур статистической обработки в статистическую СУБД и создание для них фиксированных команд и внутренних переходников, входящих в статистическую СУБД. Наименее гиб-

кий способ, т.к. для включения других статистических пакетов нужно изменять СУБД;

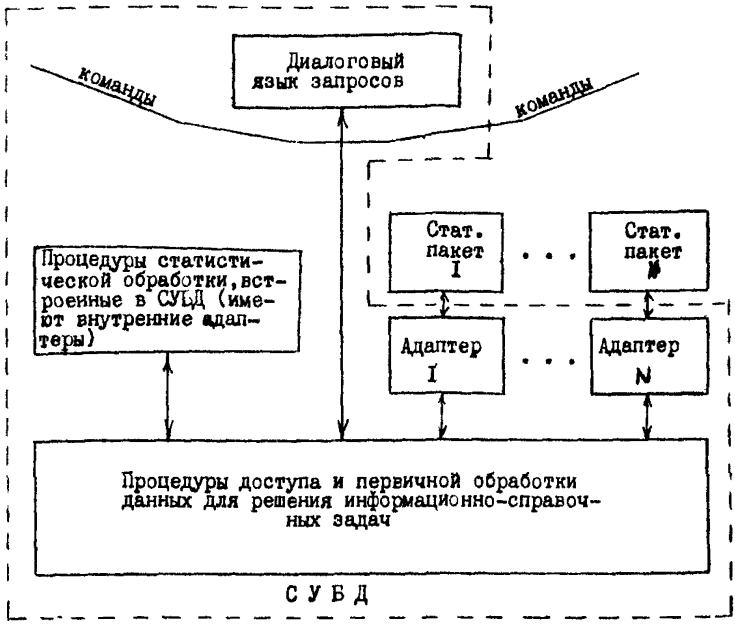
- включение статистических пакетов в БнД путем составления программ на языке запросов СУБД с указанием форматов входного набора данных статистического пакета. Более гибкий способ, т.к. позволяет не изменять СУБД, а лишь дополнять ее новыми программами;

- включение статистических пакетов в БнД путем предоставления пользователю возможности самому формировать команды для обращения к своему пакету. Пользователь фактически сам формирует переходники, передавая управляющей программе СУБД достаточную для этого информацию по сценарию диалога.

Указанные три варианта не исключают друг друга, а могут присутствовать одновременно. Например, наиболее распространенные статистические функции(суммирование, подсчет среднего и др.) могут быть жестко включены в СУБД со своими вполне определенными командами. Вместе с диалоговым языком доступа они могут обеспечить необходимый минимум статистических потребностей пользователей. В то же время в СУБ могут входить системы подключения статистических пакетов, обеспечивающих сложную обработку (аналитическое и статистическое моделирование, имитация). Это включение может идти либо по второму, либо по третьему варианту.

3.4. Структура программного и информационно-лингвистического обеспечения

Структура программного и информационно-лингвистического обеспечения статистического банка данных представляется следующей:



Данная структура реализует третий вариант включения процедур в статистический банк данных.

Этот вариант наиболее предпочтительный, но он и самый сложный в реализации. Можно развивать статистическую СУБД по этапам: сначала второй вариант и постепенный переход к третьему.

Язык запросов должен быть единым для информационно-справочного обслуживания, статистических расчетов, включать средства создания новых команд и новых функций, а также описания данных и манипулирования данными.

Возможности описания данных и их структур, включаемые в язык запросов, должны предоставлять широкий набор типов данных,

определения ключевых слов и примечаний к любому данному, оперирование с пропущенными значениями, задание информации о достоверности данных. Манипулирование данными
обеспечивает возможность:

- агрегирования;
- разбиения на подмножества и группирования;
- обработки данных по статике и динамике объекта;
- манипулирования выходными таблицами, что ограничено в обычных статистических пакетах;
- анализа представительности выборок данных;
- манипулирования метаданными.

Статистическая обработка имеет дело с таблицами, следовательно, статистические СУБД и языки запросов должны быть реляционными. В качестве языка запросов можно использовать реляционную алгебру или ее подмножество. Если же СУБД не реляционная, то могут быть разработаны программы, манипулирующие с ней как с реляционной. Например, имеющиеся БД на ИНЕС иерархической структуры можно рассматривать как реляционные, в виде таблиц. К таким таблицам может применяться реляционное исчисление.

Диалоговые языки запросов должны основываться на системе "меню", чтобы можно было только указывать, а не набирать. Затраты энергии пользователя при этом сводятся к минимуму, как и число ошибок. Для создания выходных таблиц пользователь выбирает из БД подмножество реквизитов путем указания их в командах языка запросов.

Требования к статистической СУБД.

Сформулируем основные требования к статистической СУБД, которые являются основой создания статистических банков данных:

- возможность присвоения имен, меток и комментариев данным в БД;

- проверка достоверности данных и управление пропущенными значениями;

- средства защиты данных от несанкционированного доступа и восстановление после сбоев;

- набор простых статистических процедур (гистограммы, простые линейные регрессии, графики на АЦПУ, видеограммы и др.,которые всегда будут использованы различными пользователями); возможности комбинирования их;

- возможности табулирования итоговых данных;

- интерфейс со статистическими пакетами путем создания программ на языке запросов, либо путем создания новых команд в диалоге.

3.5. Возможная схема реализации программ-адаптеров

Программы-адаптеры являются центральной частью описанных программных средств. Последовательность функционирования программных средств может быть предложена следующая (см. рис. 3.I):

- выполнение программ выборки, извлекающих требуемые данные из базы с размещением их в определенном формате в промежуточном наборе данных на МД;

- выполнение программ-адаптеров, читающих данные из промежуточного набора данных и преобразовывающих их к необходимому формату; преобразованные данные помещаются в промежуточный набор данных на МД;

- выполнение программ статистической обработки, читающих данные из промежуточного набора данных и обрабатывающих их по

определенным алгоритмам (например, статистического анализа, имитационного моделирования, оптимизационных задач), выводящим результаты обработки на АЦПУ и графопостроитель.

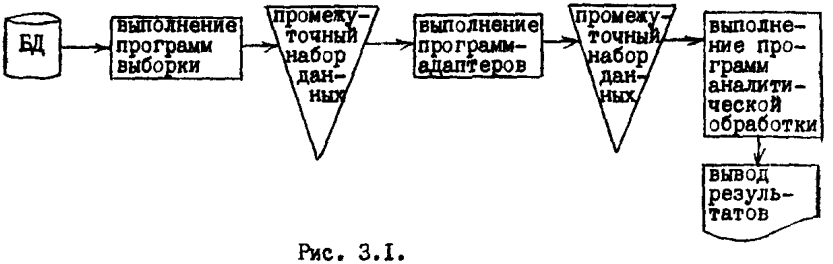


Рис. 3.I.

Выполнение указанных программных средств возможно как последовательно по шагам задания, так и в одном шаге под управлением монитора. В последнем случае возможна диалоговая организация функционирования программного комплекса с получением управляющих параметров от эксперта через дисплей и высвечиванием результата выполнения на дисплее.

Управляющие параметры, располагаемые в задании в случае пакетного режима. или принимаемые непосредственно от эксперта через дисплей в случае диалогового режима, определяют указания следующего характера:

- вид статистической обработки;
- продукция (ВКГ ОКП или группа однородной продукции);
- коды документов (ГОСТ, ОСТ, ТУ, протокол испытаний, карта уровня), показатели которых выбираются для анализа;
- множество реквизитов, выбираемых для обработки (например, обозначение и наименование марки или модели изделия, дата регистрации, код показателя, название показателя, численное значение показателя и др.).

Предлагаемый подход предусматривает многоэтапную систему прохождения реквизитов, начиная с формирования входных машинно-ориентированных форм (МОФ) и заканчивая представлением результатов их обработки в виде таблиц и графиков (см. рис. 3.2):

- входные МОФ;
- БД;
- промежуточный набор данных I;
- промежуточный набор данных 2;
- листинг (таблицы, графики).

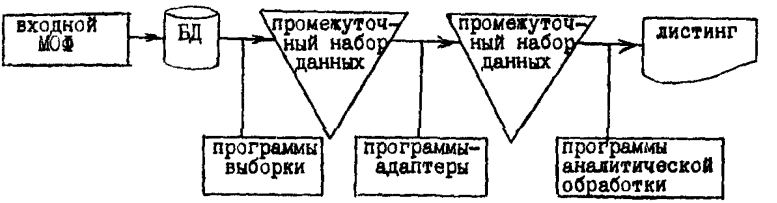


Рис. 3.2.

Во входных МОФ реквизиты представлены в виде последовательности строк документа , а в БД отображены в структурированном виде, определяемом структурой данных БД. В промежуточном наборе данных I реквизиты могут быть представлены в неформатированном и достаточно произвольном виде (например, для набора данных с последовательной организацией первые два символа каждой записи могут содержать код реквизита, остальные - его значение).

В промежуточном наборе данных 2 данные должны быть представлены в строго форматированном виде, требуемом соответствующим пакетом программ статистической обработки (как правило, набор данных имеет формат перфокарт). Листинг содержит данные, производные от начальных реквизитов, в виде таблиц и графиков.

Указанная структура программных средств, реализующая статистическую обработку данных о техническом уровне продукции, может быть предложена для БД, использующих различные СУБД. Выбор СУБД определяется многими факторами, в частности: конфигурацией имеющихся вычислительных систем (быстродействие ЭВМ, размер оперативной памяти), спецификой решаемых задач (например, объемом информации, сложностью структур данных), имеющимся опытом работы. Рекомендации по выбору СУБД не входят в задачи данной работы.

Рассмотрим возможности реализации предложенной структуры программных средств в рамках конкретной СУБД, например, системы ИНЕС, эксплуатируемой во ВНИИНМАШ.

3.6. Пример реализации на конкретной СУБД

Программы выборки могут быть реализованы:

- на языке запросов ИНЕС;
- на языке ПЛ-I, АССЕМБЛЕР, ФОРТРАН.

Язык запросов ИНЕС позволяет организовать извлечение данных из базы и некоторую их обработку перед занесением в промежуточный набор данных. Программа выборки на языке запросов может принимать управляющие параметры, передаваемые, скажем, в формате значения параметра PARM оператора EXEC (программа

%REQPARMφ) и (или) из последовательного набора данных (программа форматного чтения %INPWS). Для вывода обработанных данных в последовательный набор данных язык запросов имеет программу %A2RQPUTS.

Для реализации программ выборки на языках ПЛ-I, АССЕМБЛЕР, ФОРТРАН может быть использован табличный метод доступа (ТМД) ИНЕС. ТМД использует программу полного обхода (ППО) БД, реализованную в виде загрузочного модуля библиотеки BASE.SYSLIB и запускаемую из программы выборки на АССЕМБЛЕРЕ посредством макрокоманды INTO IPCL#EX,V или INTO ICODEX,V, а на ПЛ-I и ФОРТРАНЕ - посредством оператора CALL ICODEX. ТМД обладает тем преимуществом, что позволяет не описывать промежуточные вершины логической структуры (ДОД) БД, находящиеся на пути к нужным терминальным вершинам. Это, в частности, позволяет обеспечить независимость программ выборки от структуры базы данных.

Программа выборки на языке АССЕМБЛЕРА или ПЛ-I может использовать элементарные обращения к базе данных (спуск в подчиненную вершину, переход к вершине того же уровня, подъем на более высокий уровень, передачу значения). Элементарные средства доступа на языке АССЕМБЛЕРА реализуются посредством макрокоманд ACS1, ACS, DCONK, FCONK, RLEVEL, RTNL, ROUTE, DFR1, а на языке ПЛ-I - посредством модулей A2ACSPFM, A2MPLINA, A2NACSP, расположенных в библиотеке BASE.SRCLIB и содержащих описания необходимых функций периода компиляции и некоторых служебных переменных.

Выбор средств для реализации программы выборки может определяться рядом факторов, в частности:

- имеющимися языками программирования данного варианта ОС;
- множеством имен реквизитов, которые необходимо выбрать из базы;
- требованиями к объему памяти и времени работы программы выборки;
- сложностью структуры базы данных;
- субъективными соображениями удобства программирования на том или ином языке.

Результат работы программы выборки должен быть помещен в промежуточный набор данных, который может иметь, в частности, последовательную организацию. Практически во всех языках имеются средства как последовательного ввода, так и последовательного вывода. Промежуточный набор данных, служащий выходным для программы выборки, будет входным для программы-адаптера, поэтому к формату данных в нем высоких требований не предъявляется.

Программа-адаптер не обращается к БД, а осуществляет только обработку данных, заготовленных в промежуточном наборе, и вывод результатов обработки в другой промежуточный набор (либо в тот же самый с затиранием старой информации). Формат выходных данных программы-адаптера определяется форматом входных данных соответствующего аналитического пакета. Для каждого пакета, таким образом, может быть своя программа-адаптер, в то время как программа выборки только одна.

Поскольку программа-адаптер осуществляет лишь обработку данных без доступа к данным БД, которая может быть достаточно сложной (в зависимости от сложности входных форматов аналитического пакета), то для разработки данной программы можно

рекомендовать языки ПЛ-I, ФОРТРАН, АЛГОЛ, либо АССЕМБЛЕРА. Использование языка запросов ИНЕС едва ли оправдано, поскольку этот язык ориентирован главным образом на доступ к данным и на не слишком сложную их обработку. Так же мало подходят языки КОБОЛ и РПГ, ориентированные на большое количество операций ввода-вывода и несложную обработку данных.

После завершения работы программы-адаптера выходной набор данных, записи которого для пакетов аналитической обработки обычно имеют формат перфокарт, подается на вход соответствующего пакета. После завершения работы пакета на АЦПУ и графопостроителе можно получить машинограммы и графики, содержащие рекомендации экспертам и характеризующие положение дел в данном выбранном классе изделий. Результаты аналитической обработки предоставляют возможность интегральной оценки данных, хранимых в БД, что невозможно при обычном информационном обслуживании данными, свойственном информационным системам.

ЛИТЕРАТУРА

I. Афифи А., Эйзен С. Статистический анализ. - М., Мир; 1982.

2. Бусленко В.Н. Автоматизация имитационного моделирования сложных систем. - М.; Наука, 1977.

3. Гальперин М.В. Автоматизированная имитационная система моделирования динамики совокупности показателей технического уровня изделий машиностроения. - В сб.: Системные исследования и автоматизация в метрологическом обеспечении НИС и управлении качеством. - Львов, 1986, с. 110-112.

4. Гальперин М.В. Алгоритмическое обеспечение статистического анализа для оценки технического уровня изделий электромашиностроения. - В сб.: Вопросы совершенствования технического уровня и качества продукции машиностроения на базе стандартизации. - М.: ВНИИНМАШ, 1987, с. 60-70.

5. Левин М.Ш. Современные подходы к оценке эффективности плановых и проектных решений в машиностроении // Обзорн.информ.: Технология, оборудование, организация и экономика машиностроительного производства. Сер. 9. Автоматизированные системы проектирования и управления М.: ВНИИТЭМР, 1987-Вып. 3, с. 53.

6. Мартин Дж. Организация баз данных в вычислительных системах. - М., Мир, 1980.

7. Овчаров Л.А., Селетков С.И. Автоматизированные банки данных. - М.; Финансы и статистика, 1982.

8. РД 50-454-84. МУ САПР. Типовые математические модели объектов проектирования в машиностроении. - М.; Изд. стандартов, 1985.

ИНФОРМАЦИОННЫЕ ДАННЫЕ

РАЗРАБОТАНЫ И ВНЕСЕНЫ ВНИИНМАШ Государственного комитета СССР по стандартам.

Исполнители: к.ф.-м.н. М.В.Гальперин, А.И.Иоффин, В.Ю.Бокор.

УТВЕРЖДЕНЫ И ВВЕДЕНЫ В ДЕЙСТВИЕ Приказом ВНИИНМАШ № 375 от I декабря 1987г.

СОДЕРЖАНИЕ

Стр.

I. Технологическая схема статистического анализа и моделирования..... 4

I.I. Общие положения..... 4

I.2. Место статистического анализа и моделирования в формировании показателей технического совершенства продукции..... 6

I.3. Автоматизация моделирования как общая задача обработки данных..... 8

I.4. Технологическая схема статистического анализа и моделирования объектов стандартизации..... 8

I.5. Общая структура программного обеспечения..... I0

I.6. Понятие статистической имитации как вида моделирования..... I2

I.7. Структура алгоритмического обеспечения статистического анализа и моделирования..... I4

I.8. Алгоритм управляющего модуля..... I6

I.9. Алгоритм одного из модулей специальной статистики (расчет комплексного влияния факторов) I7

I.I0. Алгоритм статистической имитации..... I8

2. Автоматизированные системы оценки технического совершенства..... 2I

2.I. Постановка задачи создания конкретных автоматизированных программных систем..... 2I

2.2. Существующие схемы оценки технического совершенства..... 23

2.3. Технология создания диалоговой системы автоматизированного проектирования методик оцен-

..... 25

Стр.

2.4. Формальная модель автоматизированной оценки ... 29
3. Принципы построения статистических банков данных и структуры программного обеспечения 32
3.I. Постановка задачи создания статистических банков данных 32
3.2. Виды программных средств 33
3.3. Программные средства расширения возможностей СУБД 35
3.4. Структура программного и информационно-лингвистического обеспечения 37
3.5. Возможная схема реализации программ-адаптеров.. 40
3.6. Пример реализации на конкретной СУБД 43
Литература 47
Информационные данные 48

АИС-Стандарт. Технологическая схема и программное обеспечение статистического анализа и моделирования показателей технического совершенства продукции для целей автоматизации экспертизы

Рекомендации Р - 50-54-41-88

Редактор Трайнин А.И.

Мл. редактор Еремеева Т.В.

ВНИИНМАШ Госстандарта СССР

Ротапринт ВНИИНМАШ I23007 Москва, улица Шеногина, дом 4
Тираж 6I2 экз. Объем 2,I уч.-изд.л. Заказ № 673-88-I 5.03.88 г.
Цена I р.

gost.e-ecolog.ru