Рязанский государственный радиотехнический университет

Кафедра АИТП

А.С. Морозов

Моделирование систем

Конспект лекций

Рязань 2011

Введение

Моделирование как создание некой системы — системы-модели (второй системы), имеющей определенное сходство с системой-оригиналом (первой системой). Две эти реализованные системы, из которых одна рассматривается как отображение другой, связаны соотношениями подобия. Отображение одной системы в другой в этом случае является следствием выявления сложных зависимостей между двумя системами, отраженных в соотношениях подобия.

Подобие явлений означает, что данные о протекании процессов, полученные при изучении одного явления, можно распространить на все явления, подобные данным. При этом, однако, необходимо учитывать, что модель не дает и не должна давать подобия абсолютно всех процессов, содержащихся в явлении или так или иначе связанных с ним. Модель обеспечивает подобие только тех процессов, которые на ней исследуются.

Теория моделирования применяется:

а) при аналитическом отыскании зависимостей, соотношений и решений конкретных задач;

б) при обработке результатов экспериментальных исследований и испытаний различных технических устройств в тех случаях, когда результаты представлены в обобщенных «критериальных» зависимостях;

в) при создании моделей, воспроизводящих явления в установках (оригиналах), обычно больших по величине или более сложных по структуре и более дорогих, чем модели (модели анализа);

г) при создании моделей установок, позволяющих синтезировать эффективные объекты.

Классифицируя моделирование по видам и группам, целесообразно прежде всего разделить их по признакам полноты и точности воспроизведения изучаемых процессов. Последнее особенно важно, так как теория подобия и основанное на ней моделирование не могут с абсолютной полнотой воспроизводить все стороны и детали изучаемых явлений. Абсолютное подобие может мыслиться только отвлеченно и не может быть реализовано, так как это означало бы тождество, т. е. замену одного объекта или явления другим, точно таким же. Практические цели, преследуемые при решениях научных и технических задач, требуют применения моделирования в случаях, когда модель хорошо отражает изучаемый объект (оригинал) только в отношении тех явлений или входящих в эти явления процессов, которые существенны в данном исследовании, при данной постановке задачи. Таким образом, *модель — это неполная копия объекта.* Чтобы подчеркнуть это обстоятельство, иногда говорят, что «точная модель не нужна, так как при точной модели нет подобия, а есть тождество», а слишком неточная бесполезна.

Модели, которые представляют практический интерес, могут быть разделены на три (А, Б, В) способа (рис. В.2).

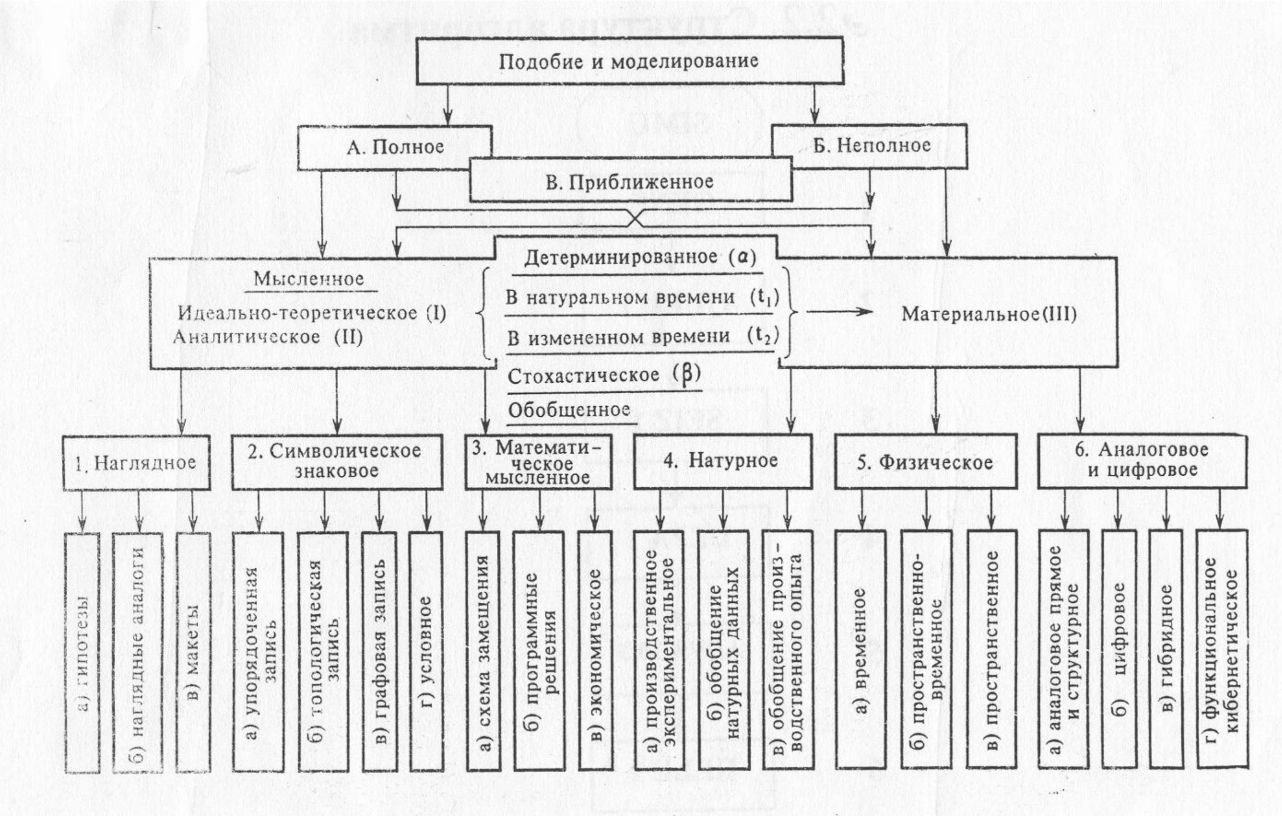


Рис. В.2. Основная классификация способов моделирования

Моделирование — это процесс создания или отыскания в природе некого объекта, замещающего исследуемый объект. Этот промежуточный объект (модель), применяемый при исследованиях, может быть реальным (материальным) объектом той же или другой природы (физическая или аналоговая модель), что и изучаемый объект (оригинал). Он может быть только мысленным объектом, воспроизводящим объект исследования с помощью логических построений, математических выкладок (математическая модель). Промежуточный объект (или квазиобъект), замещающий объект, подлежащий исследованию (оригинал), может давать надежные сведения об изучаемом объекте, если он будет его моделью. Для этого прежде всего необходимо, чтобы он был подобен изучаемому объекту. Это значит, что его параметры и параметры оригинала должны находиться в некоторых, вполне определенных соотношениях.

Из всего многообразия (рис. В.2) в дальнейшем рассмотрим только аналоговое (аналитическое), математическое и имитационное моделирование.

В инженерных задачах моделирование применяется в основном для проведения исследований на этапах проектирования, внедрения и эксплуатации сложных систем.

На этапах разработки технического и рабочего проектов систем моделирование служит для решения конкретных задач проектирования, обычно выбора оптимального по определенному критерию и при заданных ограничениях варианта системы из множества допустимых или для синтеза сложных систем.

Назначение моделирования на этапе внедрения и эксплуатации сложных систем – это проигрывание возможных ситуаций для принятия обоснованных и перспективных решений по управлению объектом. Моделирование (имитацию) также широко применяют при обучении и тренировке персонала автоматизированных систем управления, вычислительных комплексов и сетей, информационных систем в различных сферах. В этом случае моделирование носит характер деловых игр. Модель, реализуемая обычно на ЭВМ, воспроизводит поведение системы и внешней среды, а люди на основе полученных результатов в определенные моменты времени принимают решения по управлению системы.

Любую модель строят в зависимости от цели, которую ставит перед ней исследователь. Подобие процесса, протекающего в модели, реальному процессу является не целью, а условием правильного функционирования модели, и поэтому в качестве цели должна быть поставлена задача либо изучения какой-либо стороны функционирования объекта, либо определение его характеристик и т.п.

В свете изложенного, моделирование систем, как и любой производственной системы, предполагает определение состава и характеристик ее подсистем, элементов и связей между ними и внешней средой, описание процессов, протекающих в ней.

Элементы (подсистемы) можно описать тремя категориями: входы, процесс, выходы. Входы и выходы представляют собой финансовые, материальные или информационные потоки. Они имеют как непрерывный, так и дискретный характер. Процесс – это изменение состояние элементов во времени, которое может быть представлено вектором параметров, характеризующим процесс, он может меняться дискретно или непрерывно. Элементы могут иметь один и более как входов, так и выходов. Структура системы в свою очередь представляется взаимосвязанной совокупностью элементов (параметров).

Анализ и синтез технических решений систем опирается на математическое моделирование, которое позволяет решать следующие задачи:

- определение производительности всей системы и загрузки отдельных ее элементов;
- определение необходимых ресурсов производительности всех элементов системы;
- выделение наиболее существенных переменных, оценка степени влияния их изменения на исследуемые параметры системы, а также определение “узких мест”, т.е. технологических, организационных или управленческих факторов, наиболее существенно влияющих на показатели функционирования системы;
- изучение воздействий различных организационных, управленческих и технико-экономических изменений на показатели функционирования системы;
- оценка различных вариантов технических решений и стратегий управления при поиске оптимальной структуры систем.

Для моделирования систем используются как аналитическое, так и имитационное моделирование.

Аналитическое моделирование основано на косвенном описании моделируемого объекта с помощью аналитических формул. (алгебраических, интегродифференциальных, конечно-разностных, рядов и т.д.). Модель структурно не является подобной объекту моделирования и представляет формальную конструкцию, которую можно проанализировать и разрешить математическими средствами. Аналитические модели являются эффективным инструментом для решения задач оптимизации и вычисления в общем виде характеристик автоматизированных систем. Однако в ряде практических задач применение аналитических моделей затруднительно из-за их большой размерности и невозможности реализации средствами вычислительной техники. В таких случаях применяется имитационное моделирование, которое состоит в составлении вычислительного алгоритма и основано на прямом описании моделируемого объекта. При построении имитационной модели описываются законы функционирования каждого элемента объекта и связи между ними. Работа с имитационной моделью заключается в проведении имитационного эксперимента (сбор и обработка статистических данных), подобного процессу в реальном объекте. Поэтому исследования объекта на имитационной модели сводится к изучению на основе законов математической статистики характеристик процесса, протекающего в ходе эксперимента. Динамический процесс в модели протекает в так называемом системным времени, которое имитирует реальное время. Проведение имитационного моделирования часто оказывается трудоемкой и длительной процедурой и не может заменить аналитическое моделирование. Поэтому аналитическое и имитационное моделирование проводится в комплексе. Аналитическое моделирование используется для быстрого, но приближенного оценивания основных характеристик систем, а имитационное моделирование – для их уточнения.

Моделирование для исследования уже готовых объектов обычно проводится в виде организации эксперимента, обработки полученных экспериментальных данных, планирования эксперимента, позволяющего наглядно интерпретировать результаты (получить аналитическую зависимость описывающую, поведение системы при изменении параметров).

Моделирование для анализа и синтеза вновь создаваемых систем включает в себя:

1) определение параметров процессов и элементов на основе обследования и списания потоков, циркулирующих в объекте, его структуры, связей с внешней средой, а также постановки задачи и цели моделирования;

2) разработку концептуальной модели, включающую в себя разработку нескольких вариантов формализованных в общем виде и отвечающих поставленным задаче и цели моделей, выбор критерия оценки качества моделирования, выделения фрагмента объекта, подлежащего моделированию, выбор математического аппарата, описание переменных и декомпозиция модели;

3) разработку математической модели на основе выбранного математического аппарата (если выполняется имитационное моделирование, то ведется разработка моделирующего алгоритма);

4) выполнение программной реализации модели;

5) осуществление планирования проведения эксперимента и выбор средств обработки результатов эксперимента;

6) по результатам эксперимента составление заключения и выводов.

Задача курса «Моделирование систем» - дать студенту навыки по следующим вопросам:

1. Назначение, состав и основные задачи, решаемые с помощью математического моделирования.
2. Методики аналитического, имитационного моделирования, а также организации эксперимента и обработки экспериментальных данных.
3. Язык моделирования GPSS World.

Цель курса – на основе полученных знаний на практике выполнять следующее:

1. Разработать аналитические и имитационные модели производственных систем.
2. По экспериментальным данным определять функциональные зависимости между параметрами систем.
3. Выполнять корреляционно-регрессивный анализ статистических данных, полученных в результате эксперимента.
4. Выполнять планирование модельных экспериментов.
5. Реализовывать модели на языке GPSS World.

1. Моделирование как основа эксперимента

Раздел «Моделирование как основа эксперимента» предназначен для изучения методик получения для готовых объектов моделей в виде функциональных зависимостей одних параметров от других в аналитической форме. Такие зависимости на практике необходимы, во-первых, для анализа поведения объекта при определенных управлениях и условий эксплуатации, а, во-вторых, для определения возможных путей его дальнейшей модернизации. Зависимости получают путем проведения на объекте серий экспериментов с последующей обработкой полученных результатов. В разделе предлагаются методики наилучших форм представления результатов эксперимента в детерминированных и стохастических системах.

1.1. Обработка результатов эксперимента

При выборе формы представления результатов эксперимента в детерминированных системах в первую очередь рассматривается степенная зависимость, как обладающую наибольшей простотой и возможностью варьирования значений коэффициентов и показателей, характеризующих объект. Если степенная форма дает неудовлетворительные результаты, то используется подбор типа эмпирической формулы y=f(x). В стохастических системах для поиска зависимости y=f(x) используется математический аппарат корреляционно-регрессионного анализа.

Экспериментальные исследования, проводимые как в натуре, так и на моделях, должны быть предварительно тщательно продуманы не только в отношении порядка их проведения, но и в отношении выбора способов обработки результатов, оформляемых в виде связи между критериями подобия. Обработка полученных данных должна, как правило, проводиться в критериальной форме. При этом следует заметить, что термин «данные» относится ко всем материалам, накапливающимся у экспериментатора по окончании опыта. Это могут быть таблицы записей, осциллограммы, ленты регистрирующих приборов, записи на магнитных лентах, фотоснимки и т. д. Однако эта информация — только «сырые» необработанные данные. Обработанные данные составляет та же информация после выполнения над ней таких математических операций, как построение графиков, пересчет в относительные единицы, выявление функциональных зависимостей и их математическое (аналитическое) представление в виде буквенных выражений (формул). При конструировании этих выражений причинная зависимость может быть достаточно ясной, а может быть и мало заметной. В одних случаях она определяется сразу же при построении графика, а в других случаях для ее определения требуется применять статистические критерии значимости и выявлять наличие совокупности ошибок. Однако во всех случаях в результате эксперимента получают некоторую конечную выборку отсчетов из бесконечной совокупности. Чем больше выборка, тем ближе ее распределение к распределению генеральной совокупности.

При обработке эксперимента очень важно наилучшим образом выбрать форму представления его результатов.

В качестве такой формы при отсутствии других соображений следует принимать степенную зависимость $y = Ax^m U^n V^p$, как обладающую наибольшей простотой и гибкостью (возможностью варьирования значений коэффициентов и показателей для отдельных интервалов), причем следует вводить в эту степенную форму не отдельные именованные величины, а критериальные значения или, в крайнем случае, значения, представленные в относительных единицах. Метод дальнейшей конкретизации такого рода выражений, т. е. определения значений их параметров, заключается в следующем:

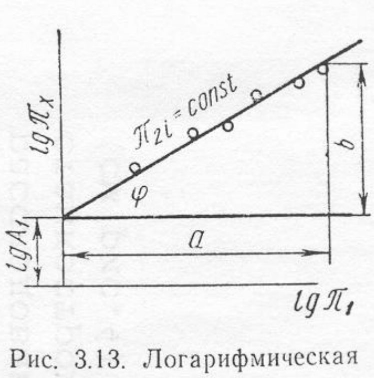


Рис. 3.13. Логарифмическая зависимость критериев подобия

1. Получив путем предварительного анализа степенную зависимость между критериями подобия $\pi_x = A\ \pi_1^m\ \pi_2^n$, необходимо так построить методику проведения экспериментов, чтобы для каждого постоянного значения определяющего критерия π_2 найти достаточно большой ряд значений неопределяющего критерия π_x в зависимости только от одного определяющего критерия π_1.

2. Полученные экспериментальные данные следует представить графически в логарифмической зависимости. Если меж-

ду критериями π_x и π_1 действительно имеет место степенная зависимость, то в логарифмической зависимости все экспериментальные точки (для каждого постоянного значения π_{2i}) укладываются на одну прямую:

$$\lg \pi_x = \lg A_1 + m \lg \pi_1,$$

где $A_1 = A\pi_{2i}^n$.

3. Далее определяется величина показателя m как тангенс угла наклона полученной прямой линии к оси абсцисс:

$$m = \operatorname{tg} \varphi = b/a;$$

где a и b находятся непосредственно масштабной линейкой.

4. Величина коэффициента A_1 определяется для ряда значений (не менее трех):

$$A_1 = \frac{1}{N} \sum_{k=1}^{N} \frac{\pi_{xk}}{\pi_{1k}^m},$$

где N — число значений π_x.

Если степенная форма критериального уравнения дает неудовлетворительные результаты, то задача конкретизации формы этого уравнения сводится к подбору типа эмпирической формулы $y=f(x)$, последующему определению значений ее параметров по методу наименьших квадратов и проверке полученной формулы. При подборе типа эмпирической формулы рекомендуется руководствоваться следующими соображениями:

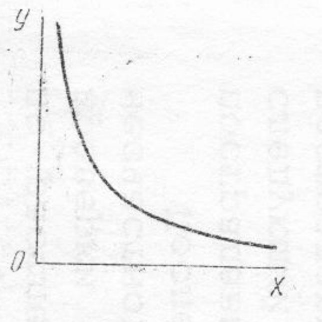


Рис. 3.14. Характерный график обратно пропорциональной связи

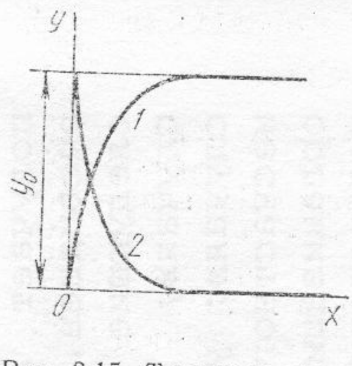


Рис. 3.15. Типичные графики экспоненциальной зависимости

1. Табличные значения переменных x и y наносят на график и по общему виду графика подбирают тип формулы.

2. Для зависимости типа обратной пропорциональности (рис. 3.14) выбирают формулу дробно-линейного вида $(A<C)$:

$$y=(Ax+B)/(Cx+D).$$

3. Если процесс является затухающим, то пробуют одну из экспоненциальных зависимостей, *1* или *2* (рис. 3.15).

4. Если процесс является периодическим (рис. 3.16), то пробуют тригонометрическую зависимость

$$y = A\sin(\alpha + 2\pi x/T),$$

или тригонометрический ряд

$$y = A_0 + A_1\cos\frac{2\pi x}{T} + A_2\cos 2\frac{2\pi x}{T} + \dots + B_1\sin\frac{2\pi x}{T} + $$
$$+ B_2\sin 2\frac{2\pi x}{T} + \dots.$$

Здесь $\pi = 3{,}14$.

5. В более общем случае, когда нельзя сделать предположений относительно ожидаемой после обработки аналитической формы, функцию представляют в виде отрезков степенного ряда полиномов той или иной степени, т. е. так называемым *уравнением регрессии*, представленным в виде ряда

$$y = b_0 + \sum b_i x_i + \sum b_{ij} x_i x_j + \sum b_{ii} x_i^2 + $$
$$+ \dots, \quad (3.29)$$

где y — исследуемый параметр; x — влияющие параметры; b_0, b_i, b_{ij}, b_{ii} — коэффициенты регрессии, которые обычно находят методом наименьших квадратов. В частном случае для получения формулы применяют многочлен

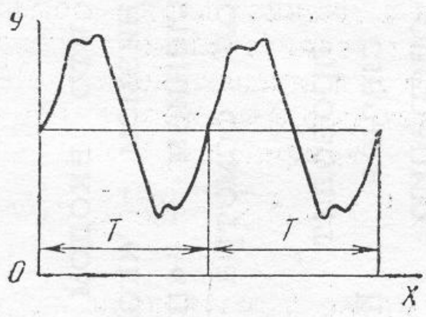


Рис. 3.16. Примерные зависимости периодического процесса с периодом T

$$y = A + Bx + Cx^2 + \dots + Mx^n.$$

При этом начинают проверку с многочлена первой или второй степени. Если она не удовлетворяется, то повышают степень многочлена. Подобрав тип эмпирической формулы, определяют численные значения ее буквенных параметров по методу наименьших квадратов.

1. Если формула нелинейна относительно параметров, то она должна быть приведена к линейному виду путем соответствующего преобразования (или разложения в ряд), в частности:

для дробно-линейной зависимости

$$Ax + B - Dy = Cxy;$$

для экспоненциальной зависимости

$$\ln\ln\frac{y_0}{y_0 - y} = m + n\ln x$$

или

$$\ln\ln\frac{y_0}{y} = m + n\ln x,$$

где $m = \ln a$; n — постоянная величина;

для тригонометрической зависимости

$$a\cos\frac{2\pi x}{T} + b\sin\frac{2\pi x}{T} = y,$$

где a=Asinά, в=Acosά .

6. В тех случаях, когда целью моделирования является определение случайной функциональной зависимости между параметрами объекта, её получают в результате специальной статистической обработки экспериментальных данных.

Допустим, по результатам эксперимента на модели необходимо найти зависимость параметра **y** от параметра **x** в виде **y = f(x).**

Для этого обработка полученных статистических данных ведётся в следующей последовательности.

1. Определяются статистические средние (оценки математических ожиданий) **y** и **x** по формулам:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i , \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i ,$$

где x_i и y_i – фактические результаты, полученные в ходе эксперимента.

2. Устанавливается функциональная зависимость между **x** и **y**.

Для дальнейшего упрощения поиска зависимости на основе корреляционно-регрессионного анализа она представляется в линейном виде $y = a_0 + a_1x$. Если искомая зависимость нелинейная, то её путём замены переменных необходимо свести к линейному виду.

<u>*Например,*</u> $y = a_0 + a_1/x$, заменить $x' = 1/x \Rightarrow y = a_0 + a_1x'$,
$y = a_0 * e^{a1x}$, заменить $y = e^{y1} \Rightarrow y_1 = \ln a_0 + a_1x$.

3. В полученной линейной зависимости определяется на сколько **x** и **y** линейно зависимы друг от друга. Оценку линейной зависимости производят с помощью <u>коэффициента корреляции R</u>

$$R = \frac{k_{x,y}}{\sigma_x * \sigma_y},$$ где $k_{x,y}$ – момент корреляции между x и y;

σ_x, σ_y - средние квадратичные отклонения по x и y.

Коэффициент R в линейной зависимости изменяется от -1 до +1. Если R отрицательный, зависимость обратно пропорциональная; если R = 0, зависимость отсутствует.

Оценка коэффициента R на основе статистических данных определяется по формуле

$$\bar{R} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} * \sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} .$$

Если $\bar{R}$ по абсолютному значению окажется меньше 0,3, то зависимость между x и y слабая ($\bar{R} \le 0{,}3$). Если окажется, что $\bar{R} \ge 0{,}7$, зависимость – сильная. В промежутке $0{,}3 < \bar{R} < 0{,}7$ средняя зависимость.

Допустим, зависимость между x и y оказалась средняя и выше. В этом случае переходят к следующему пункту.

4. Определяется количественное значение коэффициентов a_0 и a_1 по методу наименьших квадратов. Для этого определяется сумма квадратов отклонений y_i экспериментального от теоретического значения ($y_i = a_0 + a_1x_i$) по формуле

$$Z = \sum_{i=1}^{n} \left(y_i - (a_0 + a_1 x_i)\right)^2 .$$

Полученное значение Z является функцией двух коэффициентов a_0 и a_1. Их необходимо подобрать таким образом, чтобы истинное значение y_i мало отличалось от модельного значения y_i.$=a_0+ a_1x_i$. А это значит, найти такие значения a_0, a_1, при которых Z принимала бы минимальное значение. Минимум функции Z можно найти из системы уравнений с частными производными

$$\frac{\partial Z}{\partial a_0} = 0 \;; \quad \frac{\partial Z}{\partial a_1} = 0 .$$

Поиск a_0 и a_1 осуществляется следующим образом. Дифференциальные уравнения записываются на основе полученных статистических данных в виде:

$$\begin{cases} a_0 n + a_1 \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} y_i , \\ a_0 \sum\limits_{i=1}^{n} x_i + a_1 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i y_i \; . \end{cases}$$

Из системы коэффициенты определяются по формулам:

$$a_0 = \frac{\sum\limits_{i=1}^{n} y_i \sum\limits_{i=1}^{n} x_i^2 - \sum\limits_{i=1}^{n} x_i y_i \sum\limits_{i=1}^{n} x_i}{n \sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i \right)^2} ; \qquad a_1 = \frac{n \sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} y_i}{n \sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i \right)^2} .$$

.

5. После того как значения коэффициентов определены, устанавливается их значимость в функциональной зависимости, т.е. определяется насколько их значение влияет на зависимость y=f(x). Значимость коэффициентов находится путём сравнения случайных величин $t_ф$ с $t_{теор.}$, где $t_{теор}$ – теоретическая случайная величина, подчиняющаяся t-распределению и задана таблично для выбранной степени свободы V и уровня доверия α ; фактическая величина $t_ф$ для каждого из коэффициентов определяется по формуле

$$t_{aj} = \frac{a_j \sqrt{(n-2) \sum\limits_{i=1}^{n} (x_i - \bar{x})^2}}{\sqrt{\sum\limits_{i=1}^{n} (y_i - \bar{y})^2}} ,$$ j – индекс коэффициента (j = 0,1).

Если окажется, что $t_ф \le t_{теор.}$, то соответствующий коэффициент незначителен и им можно пренебречь.

6. С учетом п. 5 окончательно записывается функциональная зависимость **y = f(x).**

Пример обработки результатов эксперимента. Допустим, что по результатам планирования эксперимента имеются статистические данные по факторам x_1, x_2 и отклику y, указанные в табл. 1.

Таблица 1

Номер опыта	Значение факторов		Значение отклика
	x_1	x_2	y
1	10	5	4
2	20	15	17
3	20	5	5
4	10	15	18

Необходимо установить существует ли линейная зависимость вида $y = b_0 + b_1 x_1 + b_2 x_2$.

Для этого выполним следующую статистическую обработку данных.

1. Вычислим коэффициенты корреляции между x_1, y и x_2, y по формуле

$$\overline{R} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} * \sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}}.$$

Полученные коэффициенты $R_{x_1,y} \approx 0$, $R_{x_2,y} \approx 0{,}99$ указывают, что между x_2 и y существует сильная прямопропорциональная зависимость, а между x_1 и y – зависимость отсутствует.

2. На основании полученных результатов регрессию запишем в виде $y = b_0 + b_2 x_2$. Коэффициенты b_0, b_2 вычислим по формулам:

$$b_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \qquad (b_0 = -2);$$

$$b_2 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \qquad (b_2 = 1{,}3).$$

3. Установим значимость коэффициентов b_0 и b_2.

Для этого по формуле

$$t_{bj} = \frac{b_j \sqrt{(n-2)\sum_{i=1}^{n} (x_i - \overline{x})^2}}{\sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}}$$

определим, что фактическое значение $t_{b0} = 2{,}2$, $t_{b2} = 1{,}41$.

По таблице *t*- распределения для $\alpha = 0{,}02$ и $\upsilon = 4$ найдем $t_{теор} = 0{,}97$.

4. Так как $t_{ф} > t_{теор}$ коэффициенты b_0, b_2 значимы, окончательно искомая зависимость примет вид $y = -2 + 1{,}3x_2$.

1.2. Планирование модельных экспериментов

В случае когда на работу объекта влияет множество как внутренних, так и внешних случайных факторов функциональная искомая зависимость становится довольно сложной. Для получения таких зависимостей используется математический аппарат «планирование экспериментов». Он позволяет при ограниченном числе экспериментов получить абстрактное, но довольно точное описание функциональной зависимости в виде регрессии, состоящей из ряда всевозможных сочетаний значений внутренних и внешних факторов.

Исследователь системы, чтобы получить от модели максимум информации, должен проанализировать ее работу во всех режимах, для всех возможных сочетаний внешних и внутренних факторов и повторять каждый эксперимент множество раз. Таким образом, полученные данные, из-за большого объема, трудно поддаются анализу, да и время, затраченное на их получение чрезмерно. Для устранения указанных недостатков прибегают к так называемому «планированию» экспериментов.

Существуют два основных варианта постановки задачи планирования эксперимента:

- выбрать план, который при ограниченном числе опытов позволял бы получить наиболее достоверное значения искомой функциональной зависимости;

- выбрать план, при котором статистическая оценка функциональной зависимости может быть получена с заданной точностью при минимальном числе опытов.

Поиск плана эксперимента производится в так называемом факторном пространстве.

Факторное пространство – это множество внешних и внутренних параметров модели, значения которых исследователь может контролировать в ходе подготовки и проведения модельного эксперимента.

Поскольку факторы могут носить как количественный, так и качественный характер (например, отражать некоторую стратегию управления), значения факторов обычно называют *уровнями*. Если при проведении эксперимента исследователь может изменять уровни факторов, эксперимент называется *активным*, в противном случае – *пассивным*.

Каждый из факторов имеет верхний и нижний уровни, расположенные симметрично относительно некоторого нулевого уровня. Точка в факторном пространстве, соответствующая нулевым уровням всех факторов, называется **центром плана**, а диапазон изменения – **интервалом варьирования** фактора.

Как правило, план эксперимента строится относительно одного (основного) выходного скалярного параметра Y, который называется **наблюдаемой пе-**

ременной. Если моделирование используется как инструмент принятия решения, то в роли наблюдаемой переменной выступает показатель эффективности.

При этом предполагается, что значение наблюдаемой переменной, полученное в ходе эксперимента, складывается из двух составляющих:

$$y = f(x) + e(x),$$

где $f(x)$ – функция отклика (неслучайная функция факторов);

$e(x)$ – ошибка эксперимента (случайная величина);

x – точка в факторном пространстве (определенное сочетание уровней факторов).

Очевидно, что y является случайной переменной, так как в нее входит случайная величина $e(x)$.

Дисперсия отклика D_y, которая характеризует точность измерений, численно равна дисперсии ошибки опыта: $D_y = D_e$.

Иногда D_y называют дисперсией воспроизводимости эксперимента. Она характеризует качество эксперимента.Эксперимент называется идеальным при $D_y = 0$.

Известны различные варианты построения планов эксперимента (рандомизированный, с изменением факторов по одному, экстремальный и т. д.). Рассмотрим планирование экстремальных экспериментов.

В теории планирования эксперимента широко пользуются понятием матриц планирования эксперимента, т. е. таблицами, в которых записаны кодированные значения факторов. Каждый столбец в этой таблице (матрице планирования) называется *вектор-столбцом*. Если сумма почленных произведений любых двух вектор-столбцов матрицы равна нулю, то говорят об *ортогональности матрицы* планирования. Если точки в матрице планирования подбираются так, что точность предсказания значений параметра оптимизации одинакова на равных расстояниях от центра эксперимента и не зависит от направления, то это свойство называется *ротатабельностью*. Комбинация факторов, влияющих на проведение эксперимента, называется *уровнем факторов*. Если число факторов k известно, то можно найти число опытов N, необходимое для реализации всех возможных сочетаний уровней факторов. Это число определяется формулой $N = p^k$, где p — число уровней.

Эксперимент, при котором реализуются все возможные сочетания уровней факторов, называется *полным факторным экспериментом*. Часто применяются так называемые *дробные реплики* от полного факторного эксперимента, которыми пользуются в тех случаях, когда нужно получить линейное приближение некоторого небольшого участка поверхности отклика вместо всей поверхности. При решении такого типа задачи, например, для трех факторов можно ограничиться четырьмя опытами, если в планировании для полного факторного эксперимента типа 2^2 произведение двух влияющих факторов $x_i x_j$ приравнять третьему фактору x_{ij}.

Общая схема планирования экспериментов для решения экстремальных задач состоит из следующих этапов: 1) постановка задачи; 2) выбор параметра оптимизации; 3) выбор факторов; 4) сосоставление линейного плана; 5) реализация линейного плана и построение линейной модели; 6) поиск области экстремума; 7) описание области экстремума; 8) интерпретация результатов.

Остановимся несколько подробней на главнейших этапах.

Постановка задачи. Решение задачи начинается с ее формулировки. Исследователь должен иметь ясное, четкое, однозначное представление о цели работы. Желательно, чтобы цель исследования была сформулирована количественно, так как планирование экспериментов связано прежде всего с установлением количественных связей между входными и выходными параметрами изучаемой системы; разумеется, объект обследования должен быть управляемым.

Выбор параметра оптимизации. Одним из наиболее ответственных этапов является выбор параметра оптимизации. Он должен удовлетворять ряду требований. Желательно, чтобы параметр оптимизации был однозначным, характеризовался числами, действительно определял оптимум. Надо стремиться к тому, чтобы параметр был только один, имел ясный физический смысл и оценивался с максимальной статистической эффективностью (последнее позволяет сократить до минимума число параллельных опытов).

Простейший случай имеет место, когда заранее известен и сам параметр, и то его значение, к которому следует стремиться. При этом иногда приходится видоизменять вид параметра (например, переходить от его натуральных значений к логарифмам, обратным величинам и пр.). Если значение параметра, к которому следует стремиться, неизвестно, все же следует пытаться установить ограничения его величины хотя бы с одной стороны.

Иногда параметр оптимизации приходится изменять из-за технических трудностей, связанных, например, с отсутствием необходимых методик или достоверных методов оценки. В этих условиях можно применять параметры, дающие косвенные оценки, но поиск экстремума становится во многом интуитивным, а интерпретация результатов усложняется.

Часто возникают трудности в количественной оценке параметра оптимизации. Тогда можно использовать субъективные ранговые параметры, такие, например, как сорт, балл, класс и др. Некоторые методы планирования экспериментов вообще не требуют количественных оценок параметра оптимизации.

Выбор факторов. Не менее сложен этап выбора факторов, влияющих на изменение параметра оптимизации. Если при постановке задачи пропустить какой-нибудь сильно влияющий фактор, то вся работа окажется бесполезной. Поэтому при планировании экспериментов необходимо включать в план исследования все факторы, которые, по мнению экспериментатора, могут влиять на параметр оптимизации. Часто выбранных факторов оказывается очень много; если число их превышает 10, то возникает задача *отсеивания незначимых факторов*. Некоторые существующие в настоящее время способы отсеивания будут рассмотрены дальше. Пока же будем полагать, что факторы, влияющие на параметр оптимизации, удалось выбрать без отсеивания.

Факторы, которые по тем или иным причинам невозможно учесть в эксперименте, необходимо в течение всех опытов *стабилизировать на постоянных уровнях*.

Важным требованием, предъявляемым к факторам, является невозможность их взаимозаменяемости. Взаимозаменяемость не следует допускать даже для двух любых факторов из общей совокупности.

Выбирая факторы, рекомендуется учитывать область, ограничивающую их возможное варьирование. Желательно, чтобы факторы имели количественную оценку, хотя планирование экспериментов возможно, когда некоторые факторы представлены качественно.

После выбора факторов для каждого из них устанавливают основной уровень и интервалы варьирования. Последние следует выбирать таким образом, чтобы их величина не превышала удвоенной

ТАБЛИЦА 3.9

Уровень	Факторы (код)		
	x_1	x_2	x_3
Основной уровень (0)	0,40	840	60
Интервалы варьирования (Δx_i)	0,15	100	60
Верхний уровень (+1)	0,55	940	120
Нижний уровень (—1)	0,25	740	0

среднеквадратичной ошибки в определении данного фактора. Условия проведения опытов при изменении факторов можно представить в табл. 3.9.

Таблицы, аналогичные табл. 3.9, являются типичными, причем факторы здесь закодированы. При этом кодировании осуществляется перенос начала координат в центр (основной уровень) эксперимента и выбор масштаба с учетом интервала варьирования факторов Δx_i. Кодированные значения факторов x_i связаны с натуральными x_{i0} соотношением $x_i = (x_i \pm x_{i0})/\Delta x_i$.

Составление линейного плана. Следующий этап планирования экспериментов — составление линейного плана, реализация опытов которого преследует цель отыскания пока еще не оптимума, а лишь направления к нему.

Допустим, что в задаче варьируются только два фактора — x_1 и x_2, причем каждый на двух уровнях: +1 и —1. Все возможные комбинации факторов будут исчерпаны в четырех опытах (табл. 3.10). Линейный план, или линейная модель, характеризуется варьированием факторов на двух уровнях. Если число факторов каждого уровня равно двум, то реализуется полный факторный эксперимент типа 2^2.

ТАБЛИЦА 3.10

Номер опыта	Фактор				Кодовое обозначение строк	y
	x_0	x_1	x_2	$x_1\ x_2$		
1	+	+	+	+	aa	y_1
2	+	—	+	—	b	y_2
3	+	+	—	—	a	y_3
4	+	—	—	+	(1)	y_4

Табл. 3.10 называется *матрицей планирования полного факторного эксперимента* типа 2^2. Во втором столбце приведены значения фиктивной переменной $x_0 = +1$ (ее оценка дает величину свободного члена b_0 в уравнении регрессии); в третьем и четвертом — значения переменных x_1 и x_2 (эти два столбца и образуют собственно планирование); в пятом столбце — значения парного взаимодейст-

вия x_1x_2. Первая строка соответствует первому опыту, в котором оба фактора находятся на верхнем уровне; вторая строка — второму опыту, где фактор x_1 принимает значение нижнего уровня, а x_2 — верхнего и т д.

По результатам четырех опытов (седьмой столбец табл. 3.10) можно вычислить четыре коэффициента регрессии уравнения:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2.$$

Легко видеть, что с ростом числа факторов число опытов в полном факторном эксперименте быстро возрастает. Так, при трех факторах следует поставить $2^3 = 8$ опытов, при 5 факторах — $2^5 = 32$ опыта, а при 8 — уже $2^8 = 256$ опытов. В то же время, планируя эксперимент, исследователь может не знать заранее, в какой части изучаемой поверхности отклика находится искомый оптимум. Поэтому, начиная исследование, целесообразно получить информацию при минимальной затрате труда на проведение экспериментов, хотя, возможно, эта информация и не будет точной. В связи с этим вначале стремятся ограничиться лишь линейным описанием локальной поверхности отклика, используя дробные реплики, позволяющие упростить первые этапы исследования, сократить число опытов.

Для построения дробных реплик используют матрицы полного факторного эксперимента. Дробные реплики получают делением числа опытов соответствующего полного факторного эксперимента на число, кратное двум. Так, например, получают $^1/_2$ реплики (полуреплику), $^1/_4$ реплики (четвертьреплику), $^1/_8$ реплики и т. д.

Однако механически распределять строки матрицы на группы, например делить ее на две части, нельзя. Дробные реплики составляют заменой некоторых эффектов взаимодействия новыми независимыми переменными. Эти реплики связаны с планированием типа 2^{k-p}, где p — число линейных членов, приравненных эффектам взаимодействий. Тогда, если, например, полный факторный эксперимент типа 2^6 включает 64 опыта, то полуреплика содержит $2^{6-1} = 32$ опыта, четвертьреплика — $2^{6-2} = 16$ опытов, $^1/_8$ реплики — $2^{6-3} = 8$ опытов и т. д.

Принципиальное отличие полного факторного эксперимента от дробных реплик заключается таким образом в том, что в первом случае все линейные эффекты и взаимодействия любого порядка оцениваются независимо, а во втором случае некоторые эффекты обязательно смешаны. Число линейных эффектов, которое не смешано в данной дробной реплике, называется ее *разрешающей способностью.*

Расчет коэффициентов регрессии и определение их доверительных интервалов. По результатам опытов, проведенных в соответствии с матрицей планирования, можно подсчитать коэффициенты регрессии линейного уравнения, описывающего поверхность отклика в локальном участке вблизи выбранного основного уровня, по формуле

$$b_i = \sum_{n=1}^{N} x_{in} y_n / N,$$

где x_{in} — значение x_i в n-м опыте; y_n — значение параметра оптимизации в том же опыте.

Таким образом, способ расчета коэффициентов очень прост: столбцу y следует приписать знаки соответствующего столбца x_i, сложить все значения параметров оптимизации со своими знаками и результат разделить на число опытов матрицы планирования.

Пусть, например, в случае восьми опытов согласно полуреплике 2^{4-1} будем иметь матрицу планирования (табл. 3.14).

ТАБЛИЦА 3.14

Номер опыта	x_0	x_1	x_2	x_3	x_4	y
1	+	—	—	—	—	64
2	+	+	—	—	+	130
3	+	—	+	—	+	95
4	+	+	+	—	—	90
5	+	—	—	+	+	81
6	+	+	—	+	—	69
7	+	—	+	+	—	36
8	+	+	+	+	+	100

Тогда найдем коэффициенты b_i:

$$b_1 = \frac{(-1)64+(+1)130+(-1)95+(+1)90+(-1)81+(+1)69+(-1)36+(+1)100}{8} =$$

$$= 14,1;$$

$$b_2 = \frac{-64 - 130 + 95 + 90 - 81 - 69 + 36 + 100}{8} = -2{,}9;$$

$$b_3 = \frac{-64 - 130 - 95 - 90 + 81 + 69 + 36 + 100}{8} = -11{,}6;$$

$$b_4 = \frac{-64 + 130 + 95 - 90 + 81 - 69 - 36 + 100}{8} = 18{,}4;$$

$$b_0 = \frac{64 + 130 + 95 + 90 + 81 + 69 + 36 + 100}{8} = 83{,}1.$$

Аналогично можно было бы рассчитать и эффекты при парных взаимодействиях b_{ij} (для этого в матрицу планирования следовало бы добавить столбцы соответствующих взаимодействий).

После расчета коэффициентов регрессии следует прежде всего проверить их статистическую значимость. С этой целью рассчитывают доверительные интервалы коэффициентов регрессии (Δb_i), которые в случае планов первого порядка равны для всех коэффициентов.

В общем случае дисперсия, характеризующая ошибку в определении коэффициентов регрессии

$$S_{bi}^2 = S_{\bar{y}}^2 / Nm,$$

где $S_{\bar{y}}^2$ — дисперсия опытов: $S_{\bar{y}}^2 = \frac{1}{N}\sum_{u=1}^{N} S_u^2$, причем $S_u^2 = \frac{1}{m-1}\sum_{l=1}^{m}(y_{u,l} - \bar{y}_u)^2$ — оценка дисперсии в точках опыта; $\bar{y}_u = \sum_{l=1}^{m} y_{u,l}/m$ — среднее значение результатов; m — число повторных серий опытов.

Доверительный интервал для коэффициентов регрессии

$$\Delta b_i = \pm\, t^{\mathrm{C}} N S_{bi},$$

где t^{C} — критерий Стьюдента (может находиться по известным таблицам) *; α — уровень значимости, т. е. вероятность практически невозможных событий (обычно принимается равным 0,05 или 0,01), N — число степеней свободы.

Коэффициент регрессии можно считать статистически значимым, если его абсолютное значение равно или превышает величину доверительного интервала.

Статистический анализ уравнения регрессии. После вычисления коэффициентов регрессии и проверки их значимости проводят статистический анализ уравнения регрессии. С этой целью проверяют гипотезу об адекватности данного уравнения, т. е. ищут ответ на вопрос, соответствует ли полученное линейное уравнение изучаемому явлению или необходима более сложная модель.

Дисперсия неадекватности ($S_{\text{ад}}^2$) служит количественным показателем адекватности уравнения

$$S^2_{\text{ад}} = \sum_{n=1}^{N} (y_{n\,\text{расч}} - y_{n\,\text{эксп}})^2/(N-k-1).$$

Здесь $y_{n\,\text{расч}}$ — значение параметра оптимизации в n-м опыте, предсказанное уравнением регрессии; $y_{n\,\text{эксп}}$ — значение параметра в том же опыте, определенное экспериментально. В знаменателе выражения N — число степеней свободы при определении дисперсии неадекватности.

Схема расчета дисперсии неадекватности для рассматриваемой задачи приведена в табл. 3.15.

ТАБЛИЦА 3.15

Номер опыта	$y_{n\,\text{эксп}}$	$y_{n\,\text{расч}}$	$\lvert\Delta y\rvert = y_{n\,\text{расч}} - y_{n\,\text{эксп}}$	$\lvert\Delta y\rvert^2 = (y_{n\,\text{расч}} - y_{n\,\text{эксп}})^2$
1	64	65	1	1
2	130	130	0	0
3	95	96	1	1
4	90	87	3	9
5	81	79	2	4
6	69	70	1	1
7	36	36	0	0
8	100	101	1	1
Σ				17

Дисперсия неадекватности

$$S^2_{\text{ад}} = 17/(8-4-1) \approx 5{,}7$$

Гипотезу об адекватности проверяют с помощью критерия Фишера (F-критерия):

$$F^{расч}_{f1,\,f2} = S^2_{ад}/S^2_y,$$

где f_1 и f_2 — число степеней свободы при определении дисперсий неадекватности ($S^2_{ад}$) и опыта (S^2_y).

Гипотеза об адекватности линейной модели может быть принята, если расчетное значение F-критерия ($F^{расч}$) не превышает его табличного значения $F^{табл}$, которое имеется в специальных таблицах для выбранного уровня значимости. Для рассмотренного случая имеем

$$F^{расч}_{3;\,16} = 5{,}7/5 = 1{,}14.$$

При 5%-ном уровне значимости ($\alpha = 0{,}05$) табличное значение F-критерия $F^{табл}_{3,16} = 3{,}24$. Так как $F^{расч} < F^{табл}$, то гипотеза об адекватности линейного уравнения не отвергается, и его можно использовать для следующих этапов планирования, в частности для поиска направления движения по градиенту к оптимуму. В противном случае его нужно было бы дополнить членами высшего порядка.

Адекватность линейного уравнения можно проверить и другим способом. Свободный член уравнения регрессии b_0 является, по сути дела, оценкой результата опыта на основном уровне, когда все остальные факторы исключены. Поэтому, сделав соответствующий опыт, можно сравнить его результат с величиной свободного члена, т. е. проверить гипотезу о равенстве нулю суммы коэффициентов при квадратичных членах (нуль-гипотезу). Нуль-гипотеза может быть принята, если разность $|b_0 - y_0|$ не превышает среднеквадратичной ошибки эксперимента. Значимость этого различия иногда проверяется сопоставлением с критерием Стьюдента:

$$t^{C} \leqslant |\, b_0 - y_0 \,|\, \sqrt{N}/S_y.$$

2. Модели для параметрического анализа систем

Из всего многообразия систем рассмотрим моделирование производственных объектов. Для этого используем математический аппарат теорий массового обслуживания имитационного моделирования.

Моделирование будет направлено на описание влияния параметров на характеристики процессов (параметрическое моделирование), протекающих в системах. Вестись оно будет в той последовательности, как было предложено ранее (см. введение).

2.1. Словесное и концептуальное описание модели

Наиболее рационально строить модель функционирования системы по блочному принципу. При этом могут быть выделены три автономные группы блоков такой модели. Блоки первой группы представляют собой имитатор воздействия внешней среды *E* на систему *S*; блоки второй группы являются собственно моделью процесса функционирования исследуемой системы *S*; блоки третьей группы – вспомогательные и служат для машинной реализации блоков двух первых групп, а также для фиксации и обработки результатов моделирования.

Рассмотрим механизм перехода от описания процесса функционирования некоторой гипотетической системы к модели этого процесса. Для наглядности введем представление об описании свойств процесса функционирования системы *S*, т.е. об ее концептуальной модели M_x, как совокупности некоторых элементов, условно изображенных квадратами так, как показано на рис.2а. Эти квадраты представляют собой описание некоторых процессов исследуемого процесса функционирования системы *S*, воздействия внешней среды *E* и т.д. Переход от описания системы к ее модели в этой интерпретации сводится к исключению из рассмотрения некоторых второстепенных элементов описания, например элементов 5 – 8, 39 – 41, 43 – 47. Предполагается, что они не оказывают существенного влияния на ход процессов, исследуемых с помощью модели. Часть элементов (14, 15, 28, 29, 42) заменяется пассивными связями h_1, отражающими внутренние свойства системы (рис.2, б). Некоторая часть элементов (1 – 4, 10, 11, 24, 25) заменяется входными факторами $\boldsymbol{x}$ и воздействиями внешней среды $\boldsymbol{v_1}$. Возможны и комбинированные замены: элементы 9, 18, 19, 32, 33 заменены пассивной связью $\boldsymbol{h_2}$ и воздействием внешней среды *E*. Элементы 22, 23, 36, 37 отражают воздействие системы на внешнюю среду $\boldsymbol{y}$.

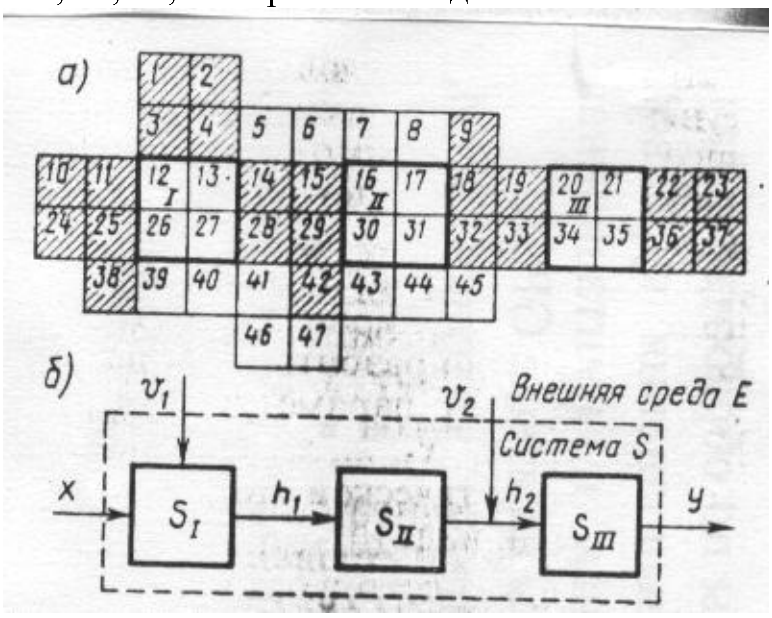


Рис.2. Модель системы: а- концептуальная, б - блочная

Оставшиеся элементы системы *S* группируются в блоки S_I, S_{II}, S_{III}, отражающие процесс функционирования исследуемой системы. Каждый из этих блоков достаточно автономен, что выражается в минимальном количестве связей между ними. Поведение этих блоков должно быть хорошо изучено и для каждого из них построена математическая модель, которая в свою очередь может содержать ряд подблоков. Построенная блочная модель процесса функционирования исследуемой системы *S* предназначена для анализа характеристик этого процесса, который может быть проведен при машинной реализации полученной модели.

После описания моделируемой системы *S* необходимо построить математическую модель процессов, протекающих в различных блоках. Математическая модель представляет собой совокупность соотношений (например, уравнений, логических условий, операторов), определяющих характеристики процесса функционирования системы *S* в зависимости от структуры системы, алгоритмов поведения, параметров системы, воздействий внешней среды *E*, начальных условия и времени.

Математическое моделирование начинается с **концептуального моделирования**, под которым в дальнейшем будем понимать переход от словесного к формальному (математическому) описание будущей модели. Концептуальная модель должна отражать общее представление об объекте в виде обобщенных математических выражений. Для ее создания необходимо выполнить следующие операции:

- определить и формализовать различными методами задачу моделирования;
- определить цель и провести выбор вариантов математической записи критерия цели моделирования (критерия эффективности);
- определить границы модели и внешней среды в свете поставленной задачи и цели;
- выбрать окончательный математический аппарат, с помощью которого можно описать поставленную задачу и цель;
- провести при необходимости детализацию модели;

Для словесного описания и разработки концептуальной модели в дальнейшем будем использовать примеры имитационного моделирования систем обслуживания.

2.1.1. Составление словесного описания объекта

На этапе словесного описания проводится подробное рассмотрение процесса функционирования объекта *S* или его фрагмента. При этом подробно описываются все элементы объекта и определяются его параметры, устанавливаются связи между ними, проводится описание внешней среды *E* в виде входных и выходных потоков или сигналов и их преобразование в ходе моделирования, т.е. описываются процессы и его параметры. Результатом описания объекта, как правило, является функциональная схема (структура объекта), где в качестве элементов фигурируют узлы модели с описанием выполняемых функций и связи между элементами в виде информационных или материальных потоков (S_I , S_{II} рис.2). В случае имитационного моделирования функциональная схема в дальнейшем будет определять структуру моделирующего алгоритма.

Например, объектом моделирования является вычислительный центр (ВЦ). Входной поток – студенты, которые приходят с интервалом времени 8 ± 2 мин. в ВЦ для работы на ЭВМ1 (ЭВМ отладки программ) и ЭВМ2. Часть входного потока (25%) поступает на ЭВМ1, а затем на ЭВМ2. Другая часть поступает непосредственно на ЭВМ2. После работы на ЭВМ2 20% студентов возвращаются для повторной работы на ЭВМ1 и ЭВМ2. На входе ВЦ имеется очередь, максимальная длина которой 4 человека. ЭВМ1 обслуживает студентов в течение 8 ± 1 мин., ЭВМ2 – в среднем 8,5 мин.

Структура блока S_{II} , отображающая процесс функционирования, имеет вид рис. 3.

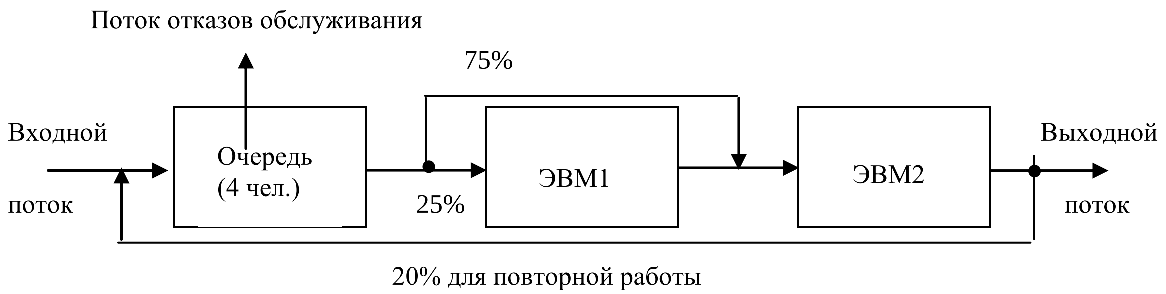


Рис. 3

ВЦ обладает следующими параметрами:

- системными: один входной случайный поток (равномерным законом распределения), две ЭВМ. ЭВМ1 обслуживает пользователей по равномерному закону, ЭВМ2 – по показательному, одна очередь на 4 студента, 20% работ после ЭВМ2 возвращаются для повторного обслуживания;

- процессорными: интервалы поступления студентов 8±2 мин., время обслуживания ЭВМ1 - 8±1 мин., ЭВМ2 – в среднем 8,5 мин.

2.1.2. Составление концептуальной модели

Описание и формализация задачи моделирования. Под задачей моделирования, как правило, понимается математическое описание структуры или процесса функционирования объекта, т.е. блоков S_I , S_{II} . Поставленная задача подвергается сначала словесному, а затем формальному описанию тех функций, с помощью которых можно реализовать структуру или процессы в объекте. Под функциями понимается преобразования входных потоков или сигналов в выходные потоки при определенных условиях и ограничениях. Для этого используются различные теории (теории потоков, теория автоматического управления, теория агрегатов, теория массового обслуживания и т.п.). Например, в системах обслуживания задачей является математическое описание процесса распределения ресурсов среди потока пользователей.

На рассматриваемом этапе в первую очередь определяется состав функций, описывающих задачу моделирования. Функции разделяются на детерминированные и стохастические. Для этих функций определяется область значения, описывается их характер. Если это случайная функция, то задаётся закон распределения и основные числовые характеристики. После определения состава функций проводится их математическая формализация таким образом, чтобы упростить реализацию этих функций и получить достаточно точные данные.

Допустим, в качестве объекта фигурирует система обслуживания, основными элементами которой являются **входной поток**, поступающий из внешней среды, **обслуживающий аппарат** и **правило обслуживания**. В этом случае основными функциями, с помощью которых можно реализовать задачу, являются функции, описывающие процесс функционирования и формализующие указанные выше элементы.

Так, по теории массового обслуживания, в имитационной модели **входной поток** можно формализовать уравнением $t_i = t_{i-1} + \tau_i$, где t_i – момент поступления в систему i-го пользователя, τ_i – интервал времени между пользователями подчиняется закону распределения $f(\tau_i)$ и обладает соответствующими числовыми характеристиками.

Обслуживающий аппарат можно формализовать уравнением в виде $t_i^{осв} = t_i^{н} + \tau_i^0$, где $t_i^{осв}$ - момент освобождения аппарата от обслуживания, $t_i^{н}$ - момент начала обслуживания, τ_i^0 - время обслуживания i-го пользователя; t_i^H может принимать два значения: $t_i^H = t_i$, если пользователь обратился в свободную систему и $t_i^H = t_{i-1}^{осв}$, если пользователь обратился в занятую систему. τ_i^0 - величина случайная, подчиняющаяся закону распределения $f(\tau_i^0)$ с соответствующими числовыми характеристиками. Если в системе есть **очереди**, то они формализуются в виде текущих очередей r_i, средних очередей $\bar{r}$, максимальных очередей $r_{\max}$ и т.п.

Правила обслуживания (распределение потоков) моделируются с помощью различных условий и сравнений. Например, по условию $t_i > t_{i-1}^{осв}$ проверяется - занята или свободна система, выбор пути движения пользователя по системе может определяться по вероятности $p \geq a$, где $0 \leq a \leq 1$, или по условию сравнения двух характеристик и т. д.

Как правило, выше рассмотренные переменные случайные и подчиняются или равномерному закону распределения в диапазоне от a до b, или показательному закону с интенсивностью λ и легко реализуются средствами имитационного моделирования.

Так, в примере с ВЦ задачей моделирования является воспроизведение процесса обслуживания студентов вычислительным центром. В этом случае основные функции блоков S_I , S_{II} запишутся в виде.

1. Входной поток: $t_i = t_{i-1} + \tau_i$, где τ_i – интервал поступления студентов в ВЦ, подчиняется равномерному закону распределения (8±2 мин).

2. Обслуживающие аппараты:

а) ЭВМ1 – $t_i^{осв1} = t_i^{н1} + \tau_i^{01}$,

б) ЭВМ2 – $t_i^{осв2} = t_i^{н2} + \tau_i^{02}$,

где τ_i^{01}, τ_i^{02} - время обслуживания, соответственно, ЭВМ1 и ЭВМ2; $t_i^{нi} = t_i$, если свободные соответствующие ЭВМ, $t_i^{нi} = t_{i-1}^{освi}$, если i-я ЭВМ занята, $t_i^{н2} = t_{i-1}^{осв1}$, если к ЭВМ2 обращаются после ЭВМ1. τ_i^{01} подчиняется равномерному закону (8±1 мин), τ_i^{02} - показательному закону (мат. ожидание 8,5 мин.).

3. Правило обслуживания:

а) отказ из-за переполнения очереди можно выполнить по условию $r_i \geq 4$, где r_i - текущая длина очереди,

б) так как 1/4 студентов обращается к ЭВМ1 остальные к ЭВМ2, тогда по вероятности Р $\geq 0,25$ можно выполнить распределение студентов, по ЭВМ,

в) повторная работа определяется по вероятности $P \leq 0,2$.

Определение цели, математическая запись критерия цели. После того, как описание задачи и ее формализация выполнены, проводится описание цели моделирования. Целью моделирования обычно является определение характеристик процессов или параметров объекта, подлежащих оценке или выбору. Как правило, целей несколько, поэтому желательно выбирать их таким образом, чтобы они могли быть реализованы одним и тем же математическим аппаратом, что и задача. Выбранные задача и цели в дальнейшем являются основополагающими и все изменения в модели, её структуре должны подчиняться только поставленной задаче и цели. В модель не должны вводиться элементы, которые не дают вклад в выполнение поставленной задачи и цели.

Для количественной оценки цели берется критерий выполнения поставленной цели, (частичная реализация блока S_{III}) . Критерием может быть характеристика объекта или оценка эффективности работы объекта. При его записи для простоты дальнейших расчетов стремятся к тому, чтобы он был единственным, а если их множество, то чтобы они были независимыми друг от друга.

Так, в системах обслуживания в качестве критерия можно выбрать:

- коэффициент загрузки обслуживающего аппарата $К_з = \sum \tau_i^0 / T$;
- среднюю длину очереди $\bar{r} = 1/n \sum_{i=1}^{n} r_i$;
- среднее время обслуживания одного пользователя $\bar{\tau}_0 = 1/n \sum_{i=1}^{n} \tau_i^0$,

где $\sum \tau_i^0$ - суммарное время обслуживания **n** поступивших за время проведения эксперимента (моделирования) Т пользователей; r_i – длина очереди, τ_i^0 - время обслуживания i-го пользователя.

Так, в выше рассматриваемом примере, целью является определение коэффициента загрузки ЭВМ1 за 8 часов (480 мин.) работы ВЦ $К_{ЗЭВМ1} = \sum \tau_i^{01}/480$.

Определение границ модели и ее переменных. На этом этапе из всего объекта выделяется фрагмент, отвечающий поставленной цели. Он подлежит дальнейшему моделированию, остальная часть объекта представляется в виде внешней среды. Если выделение фрагмента привело к изменению характеристик потоков, элементов, связей между ними, то необходима с учетом изменений дополнительная их формализация.

При описании внешней среды указываются только те связи, которые влияют на поставленные задачу и цели моделирования. Остальные связи в модели не учитываются. Обычно связь модели с внешней средой представляется в виде входных потоков (материальных, информационных) или входных сигналов. Установленные связи подлежат математическому описанию на основе ранее выбранных теорий. Как правило, входные потоки или сигналы представляются через параметры, которые фигурируют в математической записи функций модели и критерия цели. В математических записях функций и критерия цели переменные разделяются на два вида. Переменные, определяющие внешнюю среду (x, v_1, v_2), устанавливаются как независимые, неизменяемые переменные (наблюдаемые переменные), например, τ_i входного потока в системе обслуживания. Переменные, описывающие процесс функционирования объекта в модели представляются в виде управляемых, изменяемых переменных, т.е. переменных, которые в процессе моделирования можно подбирать, изменять с целью получения требуемых характеристик, например, τ_i^0 - время обслуживания. В итоге строится структура модели с расшифровкой всех внутренних и внешних связей между элементами. В примере с ВЦ согласно поставленной цели (определение коэффициента загрузки ЭВМ1) фрагмент объекта, подлежащий моделированию, будет таким как показано на рис. 2.

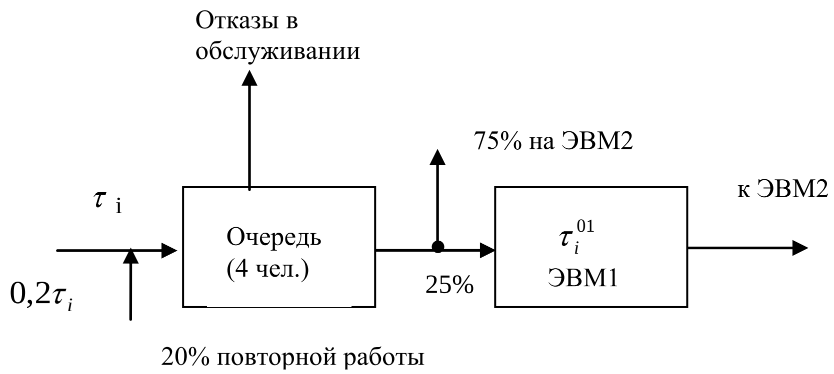


Рис.2

В модели фрагмента τ_i^{01} - управляемая переменная случайная, подчиняющаяся равномерному закону в диапазоне от 7 до 9 мин. Переменная τ_i – неуправляемая случайная с равномерным законом распределения в диапазоне от 6 до 10.

Выбор математического аппарата. После того как в общем виде составлены возможные математические записи основных функций модели и внешней среды, а также критерия цели переходят к выбору конкретного математического аппарата, с помощью которого наиболее просто и достаточно точно можно реализовать поставленные задачу и цели. Для окончательного математического описания необходимо задать исходные данные. Такая исходная информация берется из описания процесса функционирования объекта. Если этой информации недостаточно, то проводятся дополнительные исследования или выдвигаются гипотезы, которые проверяются на аналогичных объектах.

Модель объекта моделирования (система S) окончательно можно представить в виде множества величин, описывающих процесс функционирования реальной системы и образующих в общем случае следующие подмножества.

1. Совокупность входных воздействий на систему
   $x_i \in X, i = \overline{1, n_x}$.
2. Совокупность воздействий внешней среды
   $v_l \in V, l = \overline{1, n_v}$.
3. Совокупность внутренних (собственных) параметров системы
   $h_k \in H, k = \overline{1, n_H}$.
4. Совокупность выходных характеристик системы
   $y_j \in Y, j = \overline{1, n_Y}$.

При этом в перечисленных подмножествах можно выделить управляемые и неуправляемые переменные. В общем случае x_i, v_l, h_k, y_j являются элементами непересекающихся подмножеств и содержат как детерминированные, так и стохастические составляющие.

При моделировании системы S входные воздействия, воздействия внешней среды E и внутренние параметры системы являются независимыми (экзогенными) переменными, которые в векторной форме имеют соответственно вид: $\vec{x}(t) = (x_1(t), x_2(t), ..., x_{nX}(t))$, $\vec{v}(t) = (v_1(t), v_2(t), ..., v_{nV}(t))$, $\vec{h}(t) = (h_1(t), h_2(t), ..., h_{nH}(t))$, а выходные характеристики системы являются зависимыми (эндогенными) переменными и в векторной форме имеют вид $\vec{y}(t) = (y_1(t), y_2(t), ..., y_{nY}(t))$.

Процесс функционирования системы S описывается во времени оператором F_S, который в общем случае преобразует экзогенные переменные в эндогенные в соответствии с соотношением вида

$$\vec{y}(t) = F_S(\vec{x}, \vec{v}, \vec{h}, t).$$

Совокупность зависимостей основных характеристик системы от времени $y_j(t)$ для всех видов $j = \overline{1, n_Y}$ называется выходной траекторией $\vec{y}(t)$. Зависимость называется законом функционирования системы S и обозначается F_S. В общем случае закон функционирования системы F_S может быть задан в виде функции, функционала, логических условий в алгоритмической и табличной формах или в виде словесного правила соответствия.

Весьма важным для описания и исследования системы S является понятие алгоритма функционирования A_S, под которым понимается метод получения выходных характеристик с учетом выходных воздействий $\vec{x}(t)$, воздействия внешней среды $\vec{v}(t)$ и собственных параметров системы $\vec{h}(t)$. Очевидно, что один и тот же закон функционирования F_S системы S может быть реализован различными способами, т.е. с помощью множества различных алгоритмов функционирования A_S.

Как правило, математические соотношения являются описанием поведения объекта (системы) моделирования во времени *t*, т.е. отражают его динамические свойства. Поэтому математические модели такого вида принято называть динамическими моделями (системами).

Математические соотношения могут быть заданы различными способами: аналитически (с помощью формул), графически, таблично и т.д., и получены через свойства системы *S* в конкретные моменты времени, называемые состояниями, которые характеризуются векторами:

$\vec{z}'=(z_1',z_2',...,z_k')$ и $\vec{z}''=(z_1'',z_2'',...,z_k'')$,

где $z_1'=z_1(t')$, $z_2'=z_2(t')$, $\ldots$, $z_k'=z_k(t')$ в момент $t''\in(t_0,T)$; $z_1''=z_1(t'')$, $z_2''=z_2(t'')$, $\ldots$, $z_k''=z_k(t'')$ в момент $t''\in(t_0,T)$; и т.д., $k=\overline{1,n_z}$.

Если рассматривать процесс функционирования системы *S* как последовательную смену состояний $z_1(t), z_2(t), \ldots, z_k(t)$, то они могут быть интерпретированы как координаты точки в *k*-мерном фазовом пространстве, при чем каждой реализации процесса будет соответствовать некоторая фазовая траектория. Совокупность всех возможных значений состояний $\{\vec{z}\}$ называется пространством состояний объекта моделирования Z, при чем $z_k \in Z$.

Состояние системы *S* в момент времени $t_0 < t^* \le T$ полностью определяются начальными условиями $\vec{z}^0=(z_1^0,z_2^0,...,z_k^0)$, где $z_1^0=z_1(t_0)$, $z_2^0=z_2(t_0), \ldots$, $z_k^0=z_k(t_0)$, входными воздействиями $\vec{x}(t)$, внутренними параметрами $\vec{h}(t)$ и воздействиями внешней среды $\vec{v}(t)$, которые имели место за промежуток времени t^*-t_0, с помощью двух векторных управлений:

$$z(t)=\Phi(\vec{z}^0,\vec{x},\vec{v},\vec{h},t);$$
$$\vec{y}(t)=F(\vec{z},t).$$

Первое уравнение по начальному состоянию $\vec{z}^0$ и экзогенным (независимым) переменным $\vec{x}, \vec{v}, \vec{h}$ определяет вектор-функцию $\vec{z}(t)$ - эндогенные (зависимые) переменные на выходе системы $\vec{y}(t)$. Таким образом, цепочка управлений объекта «вход – состояния – выход» позволяет определить характеристики системы

$$\vec{y}(t)=F[\Phi(\vec{z}^0,\vec{x},\vec{v},\vec{h},t)].$$

В общем случае время в модели системы *S* может рассматриваться на интервале моделирования *(0, T)* как непрерывное, так и дискретное, т.е. квантованное на отрезки длинной Δt временных единиц каждый, когда $T=m\Delta t$, где $m=\overline{1,m_T}$ - число интервалов дискретизации.

Таким образом, под математической моделью объекта (реальной системы) понимают конечное подмножество переменных $\{\vec{x}(t), \vec{v}(t), \vec{h}(t)\}$ вместе с математическими связями между ними и характеристиками $\vec{y}(t)$. Если математическое описание объекта моделирования не содержит элементов случайности или они не учитываются, т.е. если можно считать, что стохастические воздействия внешней среды $\vec{v}(t)$ и стохастические внутренние параметры $\vec{h}(t)$ отсутствуют, то модель называется детерминированной в том смысле, что характеристики однозначно определяются детерминированными входными воздействиями

$$\vec{y}(t) = f(\vec{x}, t).$$

Очевидно, что детерминированная модель является частным случаем стохастической модели.

Приведенные математические соотношения представляют собой математические схемы общего вида и позволяют описать широкий класс систем. Однако в практике моделирования объектов в области системотехники и системного анализа на первоначальных этапах исследования системы рациональнее использовать типовые математические схемы: дифференцированные уравнения, конечные и вероятностные автоматы, системы массового обслуживания, сети Петри и т.д.

Не обладая такой степенью общности, как рассмотренные модели, типовые математические схемы имеют преимущества простоты и наглядности, но при существенном сужении возможностей применения. В качестве детерминированных моделей, когда при исследовании случайные факторы не учитываются, для представления систем, функционирующих в непрерывном времени, используются дифференциальные, интегральные, интегро-дифференциальные и другие уравнения, а для представления систем, функционирующих в дискретном времени, - конечные автоматы и конечно–разностные схемы. В качестве стохастических моделей (при учете случайных факторов) для представления систем с дискретным временем используются вероятностные автоматы, а для представления системы с непрерывным временем – системы массового обслуживания и т.д.

В имитационной модели функции модели можно реализовать в виде вычислительного алгоритма с помощью датчиков случайных чисел, арифметических, логических операторов, установив в них конкретные переменные и их численные значения. Для оценки критериев цели достаточно провести множество экспериментов, по результатам которых собрать и обработать статистические данные и получить функциональные зависимости.

Детализация (декомпозиция) модели. Декомпозиция модели проводится в том случае, если для выполнения поставленной цели моделирования необходимы доработки модели (более детальная проработка отдельных элементов) или же структура и состав модели излишне детализированы и тогда для выполнения поставленной цели модель можно свернут. Декомпозиция проводится также в тех случаях, если требуется уточнение или упрощение критерия цели. Если задача многокритериальная, а критерии к тому же зависят друг от друга, то её стремятся свести к единственному обобщенному критерию, остальные критерии не учитываются. Если это не удается, то путем преобразований стремятся между критериями установить линейную зависимость. Если и это не удается, подбирают удобный численный метод, с помощью которого решается многокритериальная задача с нелинейными зависимостями.

В рассмотренном выше примере декомпозицию проводить не надо. Поэтому окончательно модель можно свести к виду рис. 3.

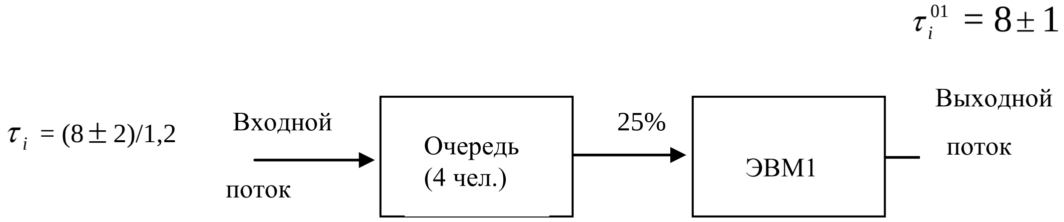


Рисунок. 3

При этом правила а, б, указанные в п. 1.2, сохраняются. На вход поступает обобщенный поток $\tau_i = (8 \pm 2)/1{,}2$.

2.2. Разработка математической модели

Исходной информацией при построении окончательных математических моделей процессов функционирования систем служат данные о назначении и условиях работы исследуемой (проектируемой) системы S, полученные в концептуальной модели. Это позволяет сформулировать требования к разрабатываемой математической модели M. При чем, уровень абстрагирования зависит от круга тех вопросов, на которые исследователь системы хотел бы получить ответ с помощью модели, т.е. от цели моделирования. Все это определяет выбор математической схемы.

Рассмотрим разработку математической модели на примере некоторых систем.

2.2.1. Аналитические модели систем обслуживания

В рассматриваемом пункте аналитическая модель представлена в виде распределения вероятностей состояния системы. Указанные распределения вероятностей для дискретного времени определяются из уравнения Маркова, описывающего так называемые Марковские процессы, а для непрерывного времени – из системы дифференциальных уравнений Колмогорова. Полученный ряд распределения вероятностей позволяет определить всевозможные характеристики объекта.

Под системой обслуживания понимается система, которая распределяет ограниченные ресурсы среди пользователей. В качестве ресурсов могут фигурировать время, материальные ценности, человеческие ресурсы, информационные ресурсы и т.д. Система включает в себя следующие элементы.

1. Входные потоки пользователей (требований, заявок). Пользователи обращаются к системе за выделением им определённых ресурсов в виде обслуживания.

2. Обслуживающие аппараты (каналы) занимаются распределением ресурсов между пользователями (обслуживанием пользователей). Обслуживаю-

щие аппараты часто имеют на своем входе так называемые очереди или средства временного хранения запросов пользователей. Время нахождения пользователя в очереди определяется пропускной способностью обслуживающего аппарата или объёмом имеющихся у него ресурсов.

3. Дисциплина обслуживания это правило, во-первых, по которому пользователь извлекается из очереди на обслуживание; во-вторых, в случае нескольких обслуживающих аппаратов как каналы привлекаются к обслуживанию, а также как регламентируется их работа; в-третьих, как оценивается качество обслуживания и распределяются потоки пользователей внутри системы.

В дальнейшем в качестве ресурса будем рассматривать время. В этом случае характеристиками элементов системы считаются:

- для входных потоков – интервал следования i-х пользователей системы τ_i;

- для обслуживающих аппаратов время обслуживания $\tau_i^{о}$. Если в системе есть очередь, то в качестве ее характеристики используется $\tau_i^{ож}$ – время нахождения i-го пользователя в очереди или используется средняя длина в очереди $\bar{r}$. В качестве общей характеристики системы используется время пребывания $\tau_i^{пр}$.

Системы, в которых проводится один акт обслуживания, называются однофазными системами. Системы, в которых проводится несколько актов обслуживания, называются многофазными.

В системах может быть несколько обслуживающих аппаратов, которые расположены функционально последовательно или параллельно.

Различают три вида систем обслуживания.

1.Система с отказом. Такая система не имеет очереди. Пользователь при обращении к занятой системе получает отказ в обслуживании.

2.Система с ограниченной очередью. В таких системах пользователь получает отказ в том случае, если очереди обслуживающих аппаратов переполнены.

3. Система без отказов (неограниченная очередь). В таких системах пользователь не получает отказ и может находиться в очереди неограниченное время.

В качестве характеристик системы, определяемых с помощью теории массового обслуживания, используется:

- для системы с отказом вероятность отказа R, вероятность обслуживания пользователя Q, иногда среднее число пользователей, получивших отказ $N_{отк}$;

- для системы с ограниченной очередью вероятность обслуживания пользователя Q, среднее время ожидания $\bar{\tau}^{ож}$, средняя длина очереди $\bar{r}$;

- для системы без отказа среднее время ожидания $\bar{\tau}^{ож}$, среднее время пребывания в системе $\bar{\tau}_{пр}$.

Основные дисциплины обслуживания.

1. Дисциплина выбора пользователя из очереди и каналов:

а) пользователь и канал выбираются в порядке очереди;

б) пользователь и канал выбираются по жребию (по вероятности, в первую очередь выбирается тот пользователь и канал, вероятность которых больше);

в) пользователь выбирается по минимально допустимому времени ожидания;

г) потоки пользователей распределяются по вероятности или по результатам сравнения значений параметров и т. д.

2. В каналах ведётся обслуживание по различным приоритетам. Приоритет может быть абсолютным и относительным. При абсолютном приоритете пользователь принимается к обслуживанию вне зависимости от состояния системы (занята или свободна). Если система занята, то обслуживание прерывается и система освобождается для обслуживания пользователя с абсолютным приоритетом. При относительном приоритете пользователь принимается к обслуживанию только после обслуживания предыдущего пользователя.

Аналитическая модель системы обслуживания включает в себя отражение потоков, обслуживающих аппаратов и правил обслуживания.

2.2.1.1. Моделирование потоков

Поток представляет собой последовательность, как правило, однородных событий. Под событием понимается обращение к системе пользователей, требований, заявок и т. д. Потоки бывают двух видов: 1) регулярные, когда интервал времени между событиями заранее определен (в детерминированных системах); 2) случайные, когда время наступления события случайно (в стохастических системах).

Одной из возможных моделей потока может быть математическая запись вероятности наступления события потока $p_i(t)$ (обращение пользователя к системе).

В простейшем случае (простейшие потоки) для определения $p_i(t)$ используется распределение Пуассона. Распределение Пуассона позволяет найти вероятность наступления k числа событий за интервал времени τ по формуле

$$p(t_{i=k}) = (\lambda\tau)^k \cdot e^{-\lambda\tau} / k!,$$

где λ - интенсивность потока.

С помощью этой формулы определяется вероятность того, что за интервал времени τ к системе будет k обращений с интенсивностью λ (число обращений в единицу времени).

Указанная формула математически определяет простейший поток. Простейшим потоком считается стационарный, ординарный поток, в котором отсутствует последействие.

Условие стационарности. Поток стационарный, если в нём вероятность появления k-событий не зависит от времени, а зависит от интервала времени τ (см. формулу Пуассона).

Условие ординарности потока. Поток считается ординарным (простым), если за интервал времени τ может наступить не более одного события. Для доказательства ординарности потока, представленного пуассоновским распределением, выполним следующие преобразования. Допустим, что за интервал времени τ в потоке не наступило события, то есть k=0, тогда

$$p_{k=0} = p_0 = e^{-\lambda\tau}.$$

Вероятность того, что за время τ наступит одно событие

$$p_{k=1} = p_1 = \lambda\tau e^{-\lambda\tau}.$$

Если сложить вероятности p_0 и p_1, получим

$$p_0 + p_1 = e^{-\lambda\tau} + \lambda\tau e^{-\lambda\tau} = (1+\lambda\tau)e^{-\lambda\tau} \approx 1.$$

Из этого следует, что за интервал времени τ в простейшем потоке может наступить не более одного события.

Отсутствие последействия в потоке. Это значит, что прошлое событие не влияет на вероятность появления последующих событий. Это условие вытекает из определения пуассоновского распределения.

Докажем что в простейшем потоке τ_i подчиняется показательному закону распределения. Функцию распределения τ можно записать в виде

$$F(t) = p(\tau \le t)$$, где t одно из значений τ.

Тогда вероятность того, что случайная величина τ не примет ни одного из значений меньше или равное t, запишется в виде

$$p(\tau > t) = 1 - F(t).$$

Из пуассоновского распределения следует, что

$$p(\tau > t) = p_0 = e^{-\lambda\tau}.$$

В этом случае

$$F(t) = 1 - e^{-\lambda\tau}.$$

Это запись функции показательного закона распределения, для которого плотность распределения записывается в виде

$f(t) = \lambda e^{-\lambda\tau}$, где $\lambda = 1/m(\tau)$ - интенсивность потока , m(τ) – математическое ожидание τ.

2.2.1.2. Моделирование обслуживающих аппаратов

В качестве математической модели обслуживающего аппарата, также как и для потоков, можно использовать вероятность состояния (обслуживающих аппаратов и очередей) в любой момент времени. Доказано, что если входные потоки простейшие, то для математического описания обслуживающих аппаратов можно использовать **теорию марковских процессов.** Под марковским процессом понимается последовательность смены состояний системы. Вероятность каждого из состояний не зависит от того, из какого состояния система перешла в искомое состояние, а определяется вероятностью настоящего состояния и вероятностью перехода. В марковском процессе фигурируют два независимых события:

1) система находится в настоящий момент в конкретном состоянии;

2) под воздействием управления (известных интенсивностей перехода) система перейдёт (или не перейдет) в другое состояние.

Следовательно, для того чтобы система перешла в следующее состояние необходимо, чтобы состоялись два выше рассмотренных события.

Допустим, под воздействием входного потока и потока обслуживания система может переходить из состояния S_0 (начальное состояние), через промежуточные состояния S_1, S_2, …, в S_n (конечное состояние). Для вероятного описания этих состояний, необходимо найти, в общем случае, вероятность перехода системы из S_i в S_j за время t ($p_{ij}(t)$). Если принять, что в момент времени t_0 (в прошлом) система находилась в состоянии S_i, а к моменту времени t (настоящее время) система находится в состоянии S_r, и под воздействием управляющих потоков (входного потока и обслуживания) в момент времени $t + \Delta t$ (будущее) система перейдет в состояние S_j , то в соответствии с формулой Маркова (модель марковского процесса) вероятность перехода системы из i-го состояния в j-ое, за время от t_0 до $t + \Delta t$ определится из уравнения

$$p_{ij}(t_0, t + \Delta t) = \sum_{r=1}^{k} p_{ir}(t_0, t)\, p_{rj}(t, t + \Delta t),$$

где $\boldsymbol{p_{ir}(t_0, t)}$ определяет вероятность перехода системы к моменту времени t из состояния S_i в промежуточное S_r или вероятность нахождения системы на момент времени t в состоянии S_r; $\mathbf{p_{rj}\ (t, t+\Delta t)}$ определяет вероятность перехода системы к моменту времени $t+\Delta t$ под воздействием управления из промежуточного состояния S_r в конечное искомое состояние S_j. На основании выше изложенного можно считать, что уравнение Маркова есть аналитическая модель системы обслуживания.

Как правило, при описании системы задаются значения вероятности начального состояния системы (выполнение первого события Марковского процесса) и вероятности (интенсивности) одношагового перехода системы из одного состояния в другое (второе событие). Поэтому, зная вероятности исходного состояния системы, можно определить вероятности состояний на последующих шагах.

Для описания динамики поведения системы обслуживания Колмогоровым на основе уравнения Маркова была предложена система дифференциальных уравнений вероятностей состояния системы. Допустим, что система находится в состоянии S_i. Известны интенсивности перехода из состояния S_i в возможные состояния S_j $\lambda_{ij}(t)$, которые полностью определяют вероятности одношагового перехода. Тогда граф примерной системы будет примет вид рис. 4.

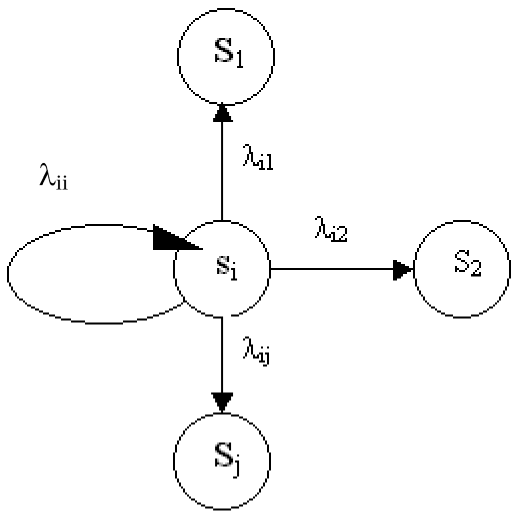


Рис. 4

Пользуясь уравнением Маркова, Колмогоров составил систему дифференциальных уравнений (аналитическую модель системы обслуживания) вида

$$dp_{j}(t)/dt = \sum_{i=1}^{k} \lambda_{ij}(t)\, p_i(t)\ .$$

Как правило, в момент времени t=0 известны: $p_i(0)$ и $\lambda_{ij}(0)$.

Решением дифференциальных уравнений являются вероятности состояний объекта, которые полностью определяют характеристики системы обслуживания.

Методика составления дифференциальных уравнений Колмогорова. Уравнения Колмогорова составляются в следующей последовательности.

1.Задаются возможные состояния системы (допустим s_1, s_2, s_3, s_4).

2.Задаются интенсивности переходов из s_i состояния в состояние s_i λ_{ij}.

3.Строится граф системы, вершинами которого являются состояния, дугами - интенсивности переходов. В рассматриваемом случае граф может иметь вид рис. 5.

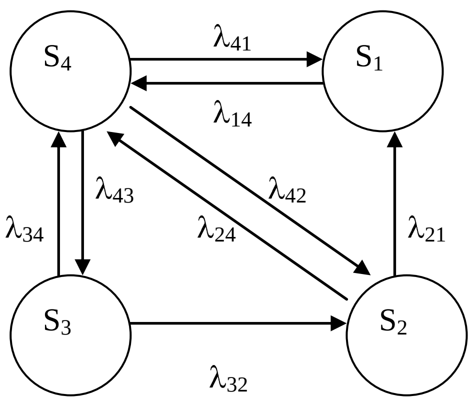


Рис. 5

4. Составляется система дифференциальных уравнений. Число уравнений соответствует числу состояний системы. Допустим, необходимо записать урав-

нение для состояния s_i. Запись ведется с использованием мнемонического правила в следующей последовательности. В левой части уравнения записывается 1-я производная вероятности состояния s_i $dp_i(t)/dt$. В правой части уравнения записывается алгебраическая сумма произведений интенсивности перехода и вероятности состояния системы. При этом произведение положительно, если система входит в описываемое состояние и отрицательно, если система выходит из этого состояния.

В рассматриваемом примере уравнение для состояния s_4 примет вид

$$dp_4(t)/dt = -\left[\lambda_{41} + \lambda_{43} + \lambda_{42}\right] p_4(t) + \lambda_{34} p_3(t) + \lambda_{24} p_2(t) + \lambda_{14} p_1(t).$$

Для того чтобы исключить нулевые решения системы, необходимо вместо одного из уравнений записать $\sum_{i=0}^{n} p_i(t) = 1$.

Если необходимо определить вероятности состояний системы в установившемся режиме ($t = \infty$), то систему дифференциальных уравнений приводят к системе линейных алгебраических уравнений.

Зная вероятности системы, как результат решения системы дифференциальных уравнений, можно определить различные характеристики системы обслуживания.

Пример решения системы в графическом виде показан на рис.6, где $t_ф$ - финальное время, при котором все вероятности системы придут к своим установившимся (финальным) значениям, а система дифференциальных уравнений сведется к системе алгебраических уравнений

$\sum_{i=0}^{k} \lambda_{ij} p_i^* = 0$, где $p_i(t_ф) = p_i^*$ - финальная вероятность.

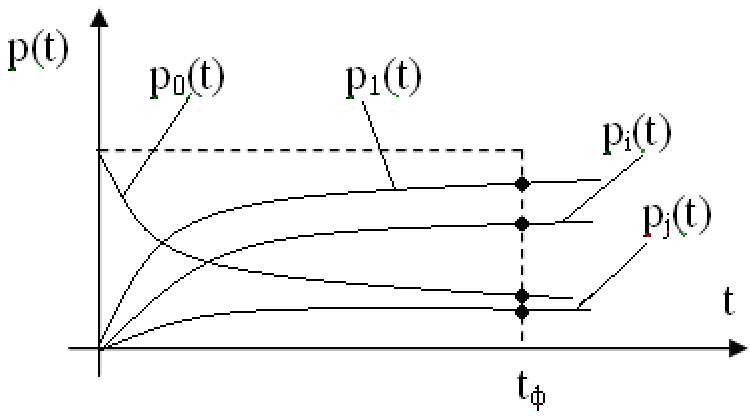


Рис. 6

2.2.1.3. Основные вероятностные характеристики систем обслуживания

1. Вероятность отказа системы в обслуживании R определяется по формуле

$$R = \sum_{i=1}^{k} p_i \ ,$$

где p_i - вероятность состояния системы, при котором пользователю даётся отказ в обслуживании.

2. Вероятность обслуживания Q=1-R .

3. Среднее время обслуживания одного пользователя

$$\bar{t}^0 = 1/\mu \ ,$$

где μ- интенсивность обслуживания (формула справедлива для простейших потоков обслуживания).

4. Средняя длина очереди на обслуживание

$$\bar{r} = \sum_{l=1}^{r} l p_l \ ,$$

где p_l – вероятность того, что в очереди находится *1* пользователей.

Средняя длина очереди математически связана со средним временем ожидания пользователя обслуживания $\bar{\tau}^{ож}$ следующим образом $\bar{r} = \bar{\tau}^{ож}\lambda$,

где λ - интенсивность обращения пользователей в систему .

5. Коэффициент загруженности i -го обслуживающего аппарата

$$\bar{k}_i = p_i \ ,$$

где p_i-вероятность занятого состояния соответствующего обслуживающего аппарата.

Пример аналитической модели системы обслуживания.

Допустим, что в примере со студенческим ВЦ, рассмотренном ранее, входной поток и обслуживание простейшие (τ_i ,τ_i^o – подчиняются показательному закону соответственно с м.о.=9,6 и с м.о.=8).

В этом случае модель ВЦ в виде системы дифференциальных управлений Колмогорова будет определена следующим образом.

1. Возможные состояния ВЦ:

S_0 – ВЦ свободен; S_1 – в ВЦ один студент и он обслуживается; S_2 - в ВЦ два студента один обслуживается, один стоит в очереди; S_3 – в ВЦ три студента один обслуживается, два стоят в очереди; аналогично определяется S_4; S_5 – ВЦ полностью занят.

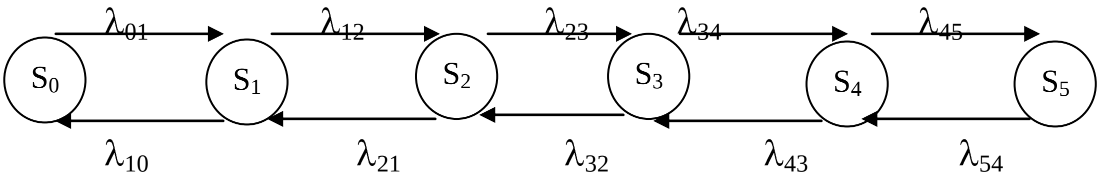
2. Интенсивности переходов.

Поскольку в ЭВМ1 на обслуживание поступает только 25% студентов (см. пример в п.1), следовательно, интенсивности, с которыми осуществляются переходы в состояния S_1, S_2 … S_5, определятся следующим образом:

1) $\lambda_{01} = \lambda_{12} = \lambda_{23} = \lambda_{34} = \lambda_{45} = \dfrac{1}{4\tau_i} = \dfrac{1}{38{,}4} = 0{,}026$ [студента в мин.];

2) $\lambda_{10} = \lambda_{21} = \lambda_{32} = \lambda_{43} = \lambda_{54}$ численно равны интенсивности обслуживания ЭВМ1 $\dfrac{1}{\tau_i^{01}} = \dfrac{1}{8} = 0{,}125$ [студента в мин.].

3. Граф ВЦ примет вид



4. Система дифференциальных уравнений (аналитическая модель ВЦ) имеет вид

$$
\begin{cases}
dP_0(t)/dt = \lambda_{10}P_1(t) - \lambda_{01}P_0(t) \\
dP_1(t)/dt = \lambda_{01}P_0(t) + \lambda_{21}P_2(t) - (\lambda_{10} + \lambda_{12})P_1(t) \\
------------------------- \\
dP_5(t)/dt = \lambda_{45}P_4(t) - \lambda_{54}P_5(t) \\
P_0(t) + P_1(t) + P_2(t) + P_3(t) + P_4(t) + P_5(t) = 1
\end{cases}
$$

Решая систему, получим значения $P_i(t)$.

5. Искомый коэффициент загрузки ЭВМ1 определится по формуле $К_3(t)=1-P_0(t)$.

2.2.2. Имитационные модели

Имитационная модель представляет собой вычислительный алгоритм, в котором реализован принцип имитации (воспроизведения) последовательности выполнения процессов, протекающих в реальном объекте, средствами вычислительной техники.

Поиск характеристик объекта осуществляется путём проведения эксперимента на модели и последующей обработки, полученных статистических данных, методами математической статистики. Имитационное моделирование используется в тех случаях, если аналитическую модель создать невозможно или же созданная модель сложна в реализации. Как правило, имитационная модель легко реализуема и имеет достаточно высокую точность.

При построении модели исследователя интересует прежде всего возможность вычисления некоторого функционала (критерия цели), заданного на множестве реализаций процесса функционирования изучаемого объекта. Например, оценка эффективности различных принципов управления системой, сравнение вариантов структуры системы, определение степени влияния изменений параметров системы и начальных условий имитации на показатель эффективности системы и т.д.

Имитационное моделирование помимо построения моделирующего алгоритма включает в себя проверку адекватности данных, полученных на модели, данным объекта, калибровку модели (изменение модели) для достижения ее адекватности, проверку на устойчивость (способность сохранять адекватность при исследовании эффективности системы во всем диапазоне значений параметров и конфигурации системы), а также оценку чувствительности модели (степени влияния изменения входных параметров на значение выходных параметров).

2.2.2.1. Принципы построения моделирующего алгоритма

В первую очередь в имитационной модели необходимо предусмотреть так называемое наращивание времени, то есть увеличение времени от начального до конечного значения. Это необходимо для того, чтобы: 1)получить временные характеристики объекта типа z_i (t); 2) получить средние по времени характеристики (статические характеристики).

Для наращивания времени в модели широко используются два принципа:

1) принцип Δt;

2) принцип особых состояний системы.

В модели время измеряется в условных единицах. Например, обслуживание в системе ведется поминутно и необходимо определить характеристику системы допустим за 8 часов работы, тогда модельное время будет составлять 480 усл. ед.

Принцип Δt. По принципу Δt время, в течение которого определяется характеристика, наращивается дискретно с шагом Δt. То есть к каждому последующему моменту добавляется Δt и определяется характеристика в новый момент времени. Для детерминированной системы z_i (t) определяется, начиная с момента t_0 и далее $t_1=t_0+\Delta t$, $t_2=t_1+\Delta t$ и т. д. В результате в искомом интервале получается дискретная функция z_i (t). Обычно известны начальные условия z_i (t_0) и интервал времени, в котором необходимо определить характеристики. Исходя из изложенного, в алгоритме необходимо предусмотреть блок наращивания времени.

Для того чтобы определить случайную временную функцию в стохастических системах по принцип Δt, в алгоритме необходимо предусмотреть: 1) блок наращивания времени; 2) имитаторы ряда значений функции в каждый из моментов времени. Определение функции в этом случае ведется следующим образом. В момент t_0 из ряда распределения вероятностей начального значения функции выбирается одно из значений. Далее наращивается время. В новый момент времени проводится серия испытаний. В результате этих испытаний создается ряд. Далее из ряда случайным образом выбирается одно из значений. Наращивается время и проводится новая серия испытаний и так далее, пока не истечет время.

Таким образом, строится одна из множеств реализаций искомой характеристики.

Для того чтобы получить статистически устойчивую характеристику, рассматривают множество реализаций (не менее 30).

Принцип Δt требует большого машинного времени. Поэтому используется только в системах, допускающих достаточно большой шаг Δt. Для устранения указанного недостатка используют принцип особых состояний.

Принцип особых состояний. Система во время функционирования может находится в двух состояниях: в особом, когда ее характеристики изменяются достаточно быстро, и не особом, когда ее характеристики или вовсе не изменяются, или изменяется незначительно. Например, в системах обслуживания особое состояние наступает с момента t_i – момента обращения к системе.

В связи с этим построение временных характеристик за время нахождения системы в не особом состоянии можно исключить, а определять характеристики на момент особых состояний, тем самым существенно сократить число экспериментов. При реализации такого принципа необходимо в алгоритме предусмотреть блок, определяющий момент вхождения системы в особое состояние t_i, и блок наращивания модельного времени $t_i = t_{i-1} + \tau_i$, где t_{i-1} – момент предыдущего особого состояния, τ_I – интервал времени между особыми состояниями.

Принцип последовательной подводки. Для построения алгоритма, имитирующего процесс функционирования объекта, часто используется **принцип последовательной проводки** пользователей системы. Алгоритм строится таким образом, чтобы пользователь последовательно проводится по его блокам, в каждом из блоков определяются его характеристики и собираются статистические данные, а затем он выводится из системы. После этого система переходит к приему следующего пользователя. Например, имеется система обслуживания, в которую поступают требования на обслуживание в случайном порядке. Система состоит из одного обслуживающего аппарата и ограниченной очереди. Дисциплина обслуживания- в порядке очереди. В этом случае принцип последовательной проводки реализуется следующим образом. Требование поступает в систему (алгоритм) (формируется t_i), затем проверяется состояние системы (занята, свободна). Если свободна, то требованию приписывается время обслуживания (имитация обслуживания), затем оно поступает в блок сбора и обработки информации и далее передается в блок формирования момента поступления нового требования входного потока. Если на момент прихода требований система занята, то оно направляется в блоки, реализующие очередь. В этих блоках требованию приписывается характеристика нахождения в очереди. Затем оно передается в блок имитирующий обслуживание. Там ему приписывается характеристика обслуживания и оно передается в блок сбора и обработки информации, а затем на вход для формирования поступления нового требования в систему и т.д. пока не истечет модельное время.

2.2.2.2. Методика построения имитационной модели

Методика построения имитационной модели (ИМ) в концептуальной части аналогична ранее рассмотренной методике моделирования. за исключением составления математической модели объекта. Из всех возможных способов организации имитационной модели рассмотрим транзактный.

В транзактном способе взаимосвязь $\boldsymbol{ФД_{ij}}$ исследуемой системы устанавливается в два этапа. Вначале ИМ представляется в виде схемы, отображающей рождение транзактов, их пространственное перемещение по схеме и уничтожение уже обслуженных транзактов. На втором этапе осуществляется программирование, при котором каждому блоку схемы ставится в соответствие определённый оператор языка моделирования, образуя программу-модель.

В программе-модели должны выполняться инициализация поступления транзакта $\boldsymbol{t_i=t_{i-1}+\tau_i}$, обслуживание в виде выделения времени $\boldsymbol{\tau_{ij}}$ для каждого $\boldsymbol{ФД_{ij}}$, перемещение по схеме модели и вывод транзакта из модели. При этом идёт проверка истечения модельного времени и условий окончания имитации. При выполнении условий окончания имитации пользователю выдаются результаты моделирования.

Имитационная модель строится в четыре этапа:

1. Выбор и составление имитаторов основных функций объекта и внешней среды (имитаторов, реализующих задачу моделирования, т.е. S_I, S_{II}).

2. Составление имитаторов сервисных функций, к которым относятся установка исходных данных, сбор и обработка статистических данных, организация эксперимента, т.е. S_{III}.

3. Составление структуры моделирующего алгоритма.

4. Описание полученного алгоритма.

Моделирование потоков и обслуживающих аппаратов моментами времени наступления события. В транзактном способе моделирование ведется следующим образом. Отсутствие последействия в потоках позволяет моменты наступления $i-$го события t_i или $t_i^{ось}$ вычислять через интервалы времени между событиями τ_i или $\tau_i^{о}$ (см. концептуальное моделирование). В этом случае момент наступления каждого последующего события будет определяться суммой момента наступления предыдущего события t_{i-1} или $t_i^{н}$ и интервала времени между этими событиями

$$t_i = t_{i-1} + \tau_i \quad \text{или} \quad t_i^{\text{îñü}} = t_i^{\text{í}} + \tau_i^{\text{î}} \ .$$

Следовательно, для моделирования потоков и обслуживающих аппаратов, в которых необходимо определить t_i или $t_i^{ось}$, достаточно знать законы распределения τ_i или $\tau_i^{о}$. Зная эти законы распределения можно легко имитировать моменты наступления события. Для этого нужно обратиться к датчику случайных чисел с установленным законом распределения, который выдает случайные числа, равные τ_i или $\tau_i^{о}$.

Выбор имитаторов основных функций. Допустим, объектом моделирования является система обслуживания, для которой основные функции - это обслуживание пользователей по определенным правилам. Система включает в себя три элемента, функционирование которых необходимо имитировать.

1.Входные потоки. Для этих потоков согласно данным концептуального моделирования необходимо выбрать датчик случайных чисел τ_i с заданным законом распределения $f(\tau)$ и числовыми характеристиками и вычислитель

$$t_{i+1}=t_i+\tau_i \ .$$

Так в рассматриваемом ранее примере с ВЦ для τ_i закон распределения равномерный в диапазоне 9,6±2,4 мин.

2. Для имитации процесса обслуживания необходим датчик случайных чисел (ДСЧ) с заданным законом распределения и числовыми характеристиками, описанными в концептуальной модели, определяющий время обслуживания τ_i^o. Помимо ДСЧ необходимы вычислители (арифметические операторы), определяющие моменты начала (t_i^H) и окончания обслуживания i-го пользователя

$$t_i^{осв}=t_i^н+\tau_i^o.$$

В рассматриваемом ранее примере закон распределения τ_i^0 равномерный в диапазоне 8±1 мин., а $t_i^H = t_i$ или $t_i^H = t_{i-1}^{осв1}$.

Если в системе имеются очереди, то для их имитации необходимы или счетчики числа поступления в очередь типа $k=k\pm1$, или ДСЧ, определяющие время нахождения пользователя в очереди $\tau_i^{ож}$, а также вычислитель момента окончания ожидания $t_i^{ож}=\tau_i^{ож}+t_i$. Датчик времени ожидания берется с законом распределения $\eta(\tau^{ож})$, который должен быть определен в концептуальной модели.

В рассматриваемом примере, в качестве имитатора очереди берется, счетчик числа студентов, находящихся в очереди (r_i).

3. Имитация правил обслуживания. В системах обслуживания помимо ранее указанных правил (переход по условию или вероятности и т.д.) существуют общие правила, по которым происходит передача данных от одного оператора к другому, а также правила присвоения значений переменным.

Так, определение состояния системы (занято, свободно) в алгоритме осуществляется по условию $t_i<t_{i-1}^{осв}$. В системах с ограниченной очередью условие того дождется ли пользователь обслуживания определяется выражением $t_i^{ож}\le t_{i-1}^{осв}$ и т. д. (см. концептуальную модель).

В рассматриваемом примере необходимо реализовать следующие правила:

- отказ в обслуживании студента, если очередь больше и равна 4(r_i GE4);
- передачу на ЭВМ1 25% студентов, остальных на ЭВМ2 (if p=0.25, ЭВМ1, ЭВМ2);
- выполнение условий $t_i\ge T$, $t_i\le t_{i-1}^{осв1}$.

Выбор имитаторов «сервисных» функций. К «сервисным» функциям относится:

Установка исходных данных. Для того чтобы каждая серия экспериментов проводилась при одинаковых и соответствующих условиям эксплуатации объекта данных, в алгоритме необходимо предусмотреть блоки присвоения начальных значений переменным, например:

$t_i=0$; $t_i^{осв}=0$; $t_i=0$; $t_i^{ож}=0$.

Сбор и обработка статистических данных. В системах обслуживания, как правило, необходимо определить случайную величину (математическое ожидание), поэтому для ее поиска используется арифметические устройства статистической обработки данных в виде вычислителя математического ожидания случайной величины

$$\overline{x} = 1/n\sum_{i=1}^{n} x_i \quad ,$$

где x_i-значение случайной величины, полученное в i-м эксперименте;

n- число проведенных экспериментов.

Помимо математического ожидания определяется дисперсия случайной величины. Для этого необходим вычислитель дисперсии

$$\sigma_{\overline{x}}^2 = 1/n(n-1)\left[n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2 \right].$$

Следовательно, для сбора и обработки статистических данных необходимы сумматоры x_i , x_i^2, вычислители средних значений случайной величины и дисперсии случайной величины.

Проведение эксперимента. Для того чтобы получить статистически устойчивые данные (с допустимой погрешностью), необходимо провести требуемое число экспериментов. Существуют три различных подхода для определения требуемого числа эксперимента.

1. Эксперимент считается законченным, если истекло модельное время Т ($t_i \geq T$). Модельное время Т обычно задается при выборе оценок цели моделирования в виде времени, за которое определяются характеристики.

2. Эксперимент заканчивается, если выполнено требуемое число прогонов, при проведении которых погрешность результатов эксперимента не превышает заданное значение ε.

3. Число экспериментов определяется автоматически. Эксперимент заканчивается, если дисперсия искомой характеристики не превышает заданное значение.

Определение требуемого числа прогонов эксперимента. Если ранее не определенно число прогонов эксперимента, то ведется специальный расчет. Для этого используется два подхода.

1. Эмпирический подход используется, если искомая случайная величина не подчиняется нормальному закону распределения. Тогда на основе теории больших чисел для получения погрешности необходимо проводить по каждой характеристике не менее 50 прогонов.

2. Если случайная характеристика подчиняется нормальному закону распределения, то используются специальные распределения (t-распределение и χ^2-распределение) для нахождения требуемого числа прогонов, обеспечивающих заданную погрешность результатов.

Допустим, определяется случайная характеристика æ с погрешностью ε.

Пусть æ- математическое ожидание искомой случайной характеристики. Из математической статистики известно, что погрешность оценки математического ожидания ε связана с t-распределением (Стьюдента) следующим соотношением

$$\varepsilon = t\sigma_x / \sqrt{N} \quad ,$$

где t-случайная величина, подчиняющаяся t-распределению, задана таблично;

σ_x- среднеквадратичное отклонение, которое обычно задано или определяется из полученного статистического ряда по формуле $\overline{\sigma}_X = \sqrt{\overline{\sigma}_X^2}$ ($\overline{\sigma_x^2}$ - оценка дисперсии x);

N- число прогонов.

Величина t определяется из таблицы по числу степеней свободы ν=N-1 и уровню доверия q=1-p/2, где p-доверительная вероятность результатов эксперимента.

Искомое число прогонов определяется по формуле: $N_{пр} \geq t^2\sigma_x^2/\varepsilon^2$.

Если в качестве æ выбирается дисперсия искомой случайной характеристики, то требуемое число прогонов вычисляется следующим образом. По известным из математической статистики формулам записывается выражение.

$$1+\varepsilon = \sqrt{X^2_{\nu,\alpha_1} / X^2_{\nu,\alpha_2}} \ ,$$

где ε - погрешность оценки дисперсии;

X^2_{ν,α_i} -случайная величина, подчиняющаяся X^2-распределению (Пирсона), задана таблично. Для выбора из таблицы значения X^2 необходимо задать число степеней свободы ν=N-1 и два уровня доверия результатов α_1=1-p/2 и α_2=1+p/2. Далее по формуле находится такая степень свободы ν, при которой

$$X^2_{\nu,\alpha_1} / X^2_{\nu,\alpha_2} = (1+\varepsilon)^2 .$$

Требуемое число прогонов, при котором дисперсия будет определена с погрешностью меньше ε, вычисляется по формуле $N_{пр} \geq \nu+1$.

Составление структуры моделирующего алгоритма. Структура алгоритма должна соответствовать функциональной схеме, рассмотренной в концептуальной модели. Моделирующий алгоритм из ранее выбранных имитаторов составляется следующим образом.

1. Составляется алгоритм, реализующий основные функции модели (задачу). В системах обслуживания это имитаторы входных потоков, имитаторы процесса обслуживания, имитаторы, определяющие правила обслуживания. Поскольку составляется имитационная модель, основой которой является воспроизведение процесса функционирования, то в первую очередь размещаются блоки, имитирующие входные потоки. Затем размещаются имитаторы процесса обслуживания (имитаторы обслуживающих аппаратов и очередей). Далее среди размещенных имитаторов устанавливают блоки или операторы, определяющие правила обслуживания.

2. В полученной структуре алгоритма размещают блоки по сбору и обработке данных, блоки по установке начальных значений переменных, блоки,

определяющие организацию проведения эксперимента (длительность эксперимента, число прогонов и т. д.)

Описание полученного алгоритма. На этом этапе проводится описание блоков, входящих в состав алгоритма, с указанием их функционального предназначения и используемых в них переменных, а также логики алгоритма и способа формирования результатов решения. Например, если это массив, то указывается его размерность и переменные. Затем проводится подробное описание связей между блоками. Если в связях возникают неоднозначные ситуации, то описываются способы выхода из них. Если необходимо уточнение какого-либо из блоков алгоритма, проводят его композицию (декомпозицию) с подробным описанием этой процедуры. Помимо этого проводится подробное описание входных и выходных данных, а также описываются способы оценки качества и точности полученных результатов и соотношений, необходимых для контроля достоверности вычислений.

Для того чтобы модель была законченным продуктом, алгоритм необходимо реализовать с помощью одного из известных языков программирования или моделирования.

Пример построения моделирующего алгоритма, основанного на принципе Δt. Составим модель производственного процесса изготовления подшипника роликового типа. Цель исследования состоит в определении оптимальных интервалов времени (t) между наладками оборудования. Для уменьшения внепроизводственных расходов типа простоя оборудования и затрат на содержание бригад наладчиков целесообразно увеличить длительность интервала времени между наладками, но при этом увеличивается доля бракованных подшипников.

Описание исследуемого процесса. Подшипник включает в себя n роликов. Качество подшипника определяется однородностью диаметров роликов z_k. Желательно, чтобы все n роликов одного подшипника были по возможности одинакового диаметра.

Подшипник считается годным, если разброс диаметров, обозначенный U, не превышает заданной величины δ. Разброс диаметров не остаётся постоянным и зависит от продолжительности работы оборудования после очередной наладки.

Целью моделирования является получение зависимости доли брака от времени между наладками q = q(t). Эта зависимость случайная функция. Моделирующий алгоритм целесообразно строить по принципу Δt, так как в системе отсутствуют особые состояния.

Построение моделирующего алгоритма. Выбор имитаторов ведется в следующей последовательности.

1. Выбор имитаторов основных функций.

Так как целью моделирования является определение доли брака q = q(t), задается временной интервал от 0 до T_{max}. Следовательно, в алгоритме должен быть блок наращивания времени от 0 до T_{max}. Основной функцией изготовления подшипника является выпуск роликов диаметром z_k. Следовательно, необходим имитатор z_k , подчиняющийся определённому закону распределения. Помимо формирователя z_k необходимы имитаторы проверки разброса диаметров роликов. Поэтому необходимо предусмотреть в алгоритме блоки, реализующие следующие вычисления:

- $U = Z_{max} – Z_{min}$, где Z_{max} – максимальный диаметр из выборки объёмом n,
- Z_{min} – минимальный диаметр из выборки объёмом n;
- $Z_{max} = \max_{n} \{Z_1, Z_2,\ldots, Z_n\}$;
- $Z_{min} = \min_{n} \{Z_1, Z_2,\ldots, Z_n\}$;
- блок проверки годности подшипника $U \le \delta$.

Имитаторы основных функций:

$Ф_1$ => Z_k (датчик случайных чисел, имитатор изготовления роликов);

$А_2$ => k < n (проверка k < n, где k-число изготовленных роликов, n-общее число роликов, входящих в подшипник);

$К_3$ => k=k+1 (счётчик числа изготовленных роликов).

$А_4$ => Z_{max} (вычислитель Z_{max});

$А_5$ => Z_{min} (вычислитель Z_{min});

$А_6$ => $U = Z_{max} – Z_{min}$ (вычислитель разброса);

$А_7$ => $U \le \delta$ (проверка подшипника на годность),

В имитаторах Ф – формирователь случайных чисел, А – арифметический оператор, К – счетчик числа событий, F – оператор присвоения.

2. Выбор имитаторов «сервисных» функций.

Долю бракованных подшипников или вероятность брака определяется по формуле $q = \frac{\overline{m}}{M}$, где $\overline{m}$ - число бракованных подшипников; М – общее число подшипников.

Следовательно, для вычисления доли брака необходимы:

- счётчик числа годных подшипников $m = \sum_{k=1}^{M} \mu_k$, где μ_k принимает значение «0», если подшипник бракован, и «1», если подшипник годен;
- вычислитель числа бракованных подшипников $\overline{m} = M - m$;
- вычислитель доли брака $q = \frac{\overline{m}}{M}$.

Имитаторы сбора и обработки статистических данных:

F_8 => $\mu_к$=1 (оператор присвоения, т.е. фиксатор годности);

F_9 => μ_0=0 (фиксатор бракованного подшипника);

K_{10} => $m = m + \mu_к$ (вычислитель числа годных подшипников);

A_{11} => $q = \frac{M - m}{M}$ (вычислитель доли брака).

Для организации проведения эксперимента в алгоритме необходимо предусмотреть следующие блоки проверки:

- $T_{l+1} \le T_{max}$.- текущее время достигло ли выбранного максимального значения ;
- счетчик числа экспериментов V;

- $V \leq M$- число экспериментов V достигло ли выбранного значения M.

Имитаторы проведения эксперимента:

A_{12} => $T_{l+1} = T_l + \tau$ (блок наращивания времени, где $\tau = \Delta t$);

A_{13} => $T_l \leq T_{max}$ (устройство проверки условия);

K_{14} => $V = V + 1$ (счётчик числа сформированных подшипников);

A_{15} => $V < M$ (оператор проверки условия).

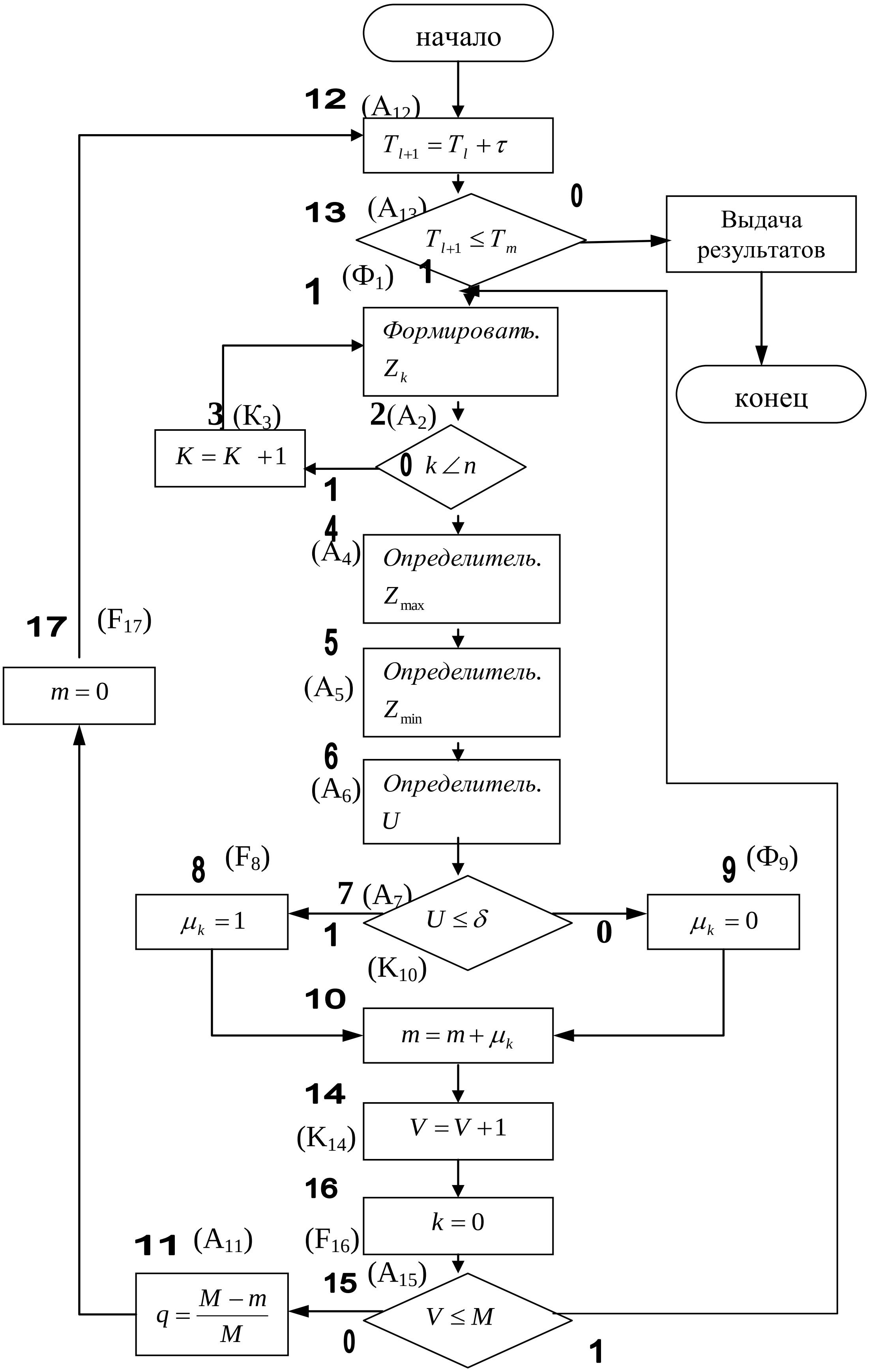
Для установки исходных данных в алгоритме выбираются два блока к = 0 (при переходе к изготовлению очередного подшипника), m = 0 (при переходе сбора данных для очередного значения T_l).

Имитаторы установки исходных данных:

F_{16} => k = 0 (оператор присвоения k = 0);

F_{17} => m = 0 (оператор присвоения m = 0).

Структура моделирующего алгоритма



Описание алгоритма. Составленный алгоритм позволяет на основе статистических испытаний определить зависимость доли брака от времени между наладками оборудования. Для получения статистически устойчивой характеристики для каждого интервала времени T_l, проводится M число испытаний.

Работает модель следующим образом. При запуске программы устанавливаются значения: $T_l = 0$, $k = 0$, $V = 0$, $\mu_k = 0$, τ, T_{max}, M, δ.

Затем определяется текущее время T_{l+1} оператором ***12***. Далее управление передаётся оператору ***13***, который проверяет, достигло ли текущее время максимально допустимого значения. Если «нет», управление передаётся формирователю роликов. Формирование роликов проводится имитатором диаметров роликов z_k. Далее управление передаётся оператору ***2***, проверяющему, все ли ролики подшипника сформированы. Если не все управление передаётся счётчику числа роликов (оператор ***3***) и далее к оператору ***1*** для формирования очередного ролика. Если сформированы все n роликов, т.е. создан массив z_k объёмом n, управление передаётся оператору ***4***, определяющему z_{max}. Далее оператором ***5*** определяется z_{min}, а оператором ***6*** разброс U. Оператором ***7*** проводится контроль годности подшипника. Если подшипник годен, управление передаётся оператору ***8***, а затем оператору ***10*** для подсчета числа годных подшипников. Если подшипник не годен, управление передаётся оператору ***9*** и далее оператору ***10***. От оператора ***10*** управление передаётся оператору ***14*** для подсчета числа проверенных подшипников. Далее оператором ***16*** путём присвоения $k = 0$ систему подготавливают к контролю следующего подшипника. Далее управление передаётся оператору ***15***, проверяющему, все ли подшипники просмотрены. Если не все, то управление передаётся оператору ***1*** для формирования очередного подшипника. Если все подшипники проконтролированы, управление передаётся оператору ***11*** для подсчета доли отказа при времени между наладками T_{l+1}. Далее оператор ***17***, подготавливает систему к вычислению доли отказов на следующем интервале времени. Если условие оператора ***13*** не выполняется (время истекло), выдаются результаты в виде зависимости $q = q(T_l)$.

Пример составления имитационной модели по принципу особых состояний. В качестве примера рассмотрим работу оператора банка. Целью моделирования выберем следующее: определение среднего времени обслуживания одного клиента за интервал времени от 0 до T, коэффициент загрузки оператора (пропускную способность) и среднее время ожидания клиента своего обслуживания.

Система оператор – клиент работает следующим образом. Оператор обслуживает клиентов в течение времени, подчиняющемуся закону $\zeta(\tau^0)$. К оператору через промежутки времени, подчиняющемуся закону $f(\tau)$, обращаются клиенты. Если оператор занят, клиент становится в очередь. Максимально допустимое время нахождения клиента в очереди $\tau_{\max}^{ож}$ подчиняется закону $\Psi(\tau_{\max}^{ож})$.

Задачей моделирования является воспроизведение процесса обслуживания клиента оператором. Цель моделирования - найти $\bar{\tau}^{o}$ - среднее время обслуживания одного клиента, $\bar{\tau}^{ож}$ - среднее время нахождения клиента в очереди и δ - пропускную способность оператора.

Выбор имитаторов основных функций. Так как рассматриваемая система является системой обслуживания, выбираем имитаторы следующих элементов.

1. Входной поток.

Для формирователя τ_i, подчиняющегося закону $f(\tau)$, выбираем датчик случайных чисел. Для определения t_i выбираем сумматор $t_i = t_{i-1} + \tau_i$.

2. Имитаторы обслуживания.

В качестве формирования τ_i^0 выбираем датчик случайных чисел, подчиняющихся закону $\zeta(\tau^0)$. Для определения $t_i^{осв.}$ выбираем сумматор $t_i^{осв.} = t_i^{н} + \tau_i$. Для определения $t_i^{н}$ выбираем два блока присвоения: $t_i^{н} = t_i$, если клиент обслуживается, не заходя в очередь; $t_i^{н} = t_{i-1}^{осв}$ - клиент обратился к занятому оператору, и был поставлен в очередь.

3. Имитаторы очереди.

Для формирования $\tau_{\max}^{ож.}$ выберем датчик случайных чисел, подчиняющихся закону $\Psi(\tau_{\max}^{ож.})$. Для определения момента окончания ожидания $t_i^{ож.}$ выбираем сумматор $t_i^{ож.} = t_i + \tau_{i\max}^{ож.}$.

4. Имитаторы правила обслуживания.

- ($t_i^{осв.}$>T) клиент не обслуживается, если момент окончания обслуживания $t_i^{осв.}$ выходит за рамки отведённого времени Т.

- $t_i < t_{i-1}^{осв.}$ клиент, поступивший к занятому оператору, становится в очередь.

- $t_i^{ож.} \le t_{i-1}^{осв.}$ клиент, у которого время нахождения в очереди ограничено и меньше времени окончания обслуживания предыдущего клиента, не обслуживается.

В итоге имитаторы основных функций:

***A*₃**: $t_i < t_{i-1}^{осв.}$; ***Ф*₄**: $\tau_{\max}^{ож} \rightarrow \Psi(\tau_{\max}^{ож.})$; ***A*₅**: $t_i^{ож.} = t_i + \tau_{i\max}^{ож.}$; ***A*₆**: $t_i^{ож.} < t_{i-1}^{осв.}$;

***F*₇**: $t_i^{н} = \tau_{i-1}^{осв.}$; ***F*₈**: $t_i^{н} = t_i$; ***Ф*₉**: $\tau_i \rightarrow \Psi(\tau_i^{o})$; ***A*₁₀**: $t_i^{осв.} = t_i^{н} + \tau_i$;

***A*₁₁**: $t_i^{осв.} \le T$; ***Ф*₁**: $\tau_i \rightarrow f(\tau_i)$, $t_i = t_{i-1} + \tau_i$.

Выбор имитаторов «сервисных» функций. 1. Имитаторы сбора и обработки данных.

Для определения указанных выше характеристик необходимы следующие блоки:

- счётчик числа обслуженных клиентов m;
- счётчик числа не обслуженных клиентов $\overline{m}$;
- сумматор, определяющий общее время обслуживания всех клиентов за интервал времени от 0 до Т: $\tilde{\tau}^0 = \sum \tau_i^0$;
- сумматор для вычисления общего времени ожидания клиентами своего обслуживания $\tilde{\tau}^{ож.}$;
- вычислитель $\tilde{\tau}^{ож.} = t_i^н - t_i$ (вычислитель фактического времени ожидания);
- вычислители: $\delta = \dfrac{m}{m + \overline{m}}$; $\overline{\tau}^{ож.} = \dfrac{\tilde{\tau}^{ож.}}{m}$; $\overline{\tau}^0 = \dfrac{\tilde{\tau}^0}{m}$.

В итоге имитаторы сбора и обработки:

К₁₂: $m = m + 1$; ***А₁₃***: $\tau_i^{ож.} = t_i^н - t_i$; ***К₁₄***: $\overline{m} = \overline{m} + 1$; ***А₁₈***: $\tilde{\tau}^0 = \sum \tau_i^0$;

А₁₉: $\tilde{\tau}^{ож.} = \sum \tau_i^{ож.}$; ***А₂₀***: $\delta = \dfrac{m}{m + \overline{m}}$, $\overline{\tau}^{ож.} = \dfrac{\tilde{\tau}^{ож.}}{m}$, $\overline{\tau}^0 = \dfrac{\tilde{\tau}^0}{m}$.

2. Имитаторы организации эксперимента.

Для организации эксперимента необходимы:

- блок окончания прогона (прогон считается законченным, если $t_i \geq T$);

- блок наращивания числа прогонов (если за время Т собрано недостаточное число данных (меньше 30), в алгоритм вводятся дополнительные блоки, с помощью которых организуются повторные прогоны модели; это блоки: счётчик числа прогонов N = N + 1 и арифметическое устройство, проверяющее условие $N \leq N_0$, где N_0 – требуемое число прогонов).

В итоге имитаторы организации испытаний:

A_2: $t_i < T$; *K_{15}*: N = N + 1; ***A_{16}***: $N \leq N_0$.

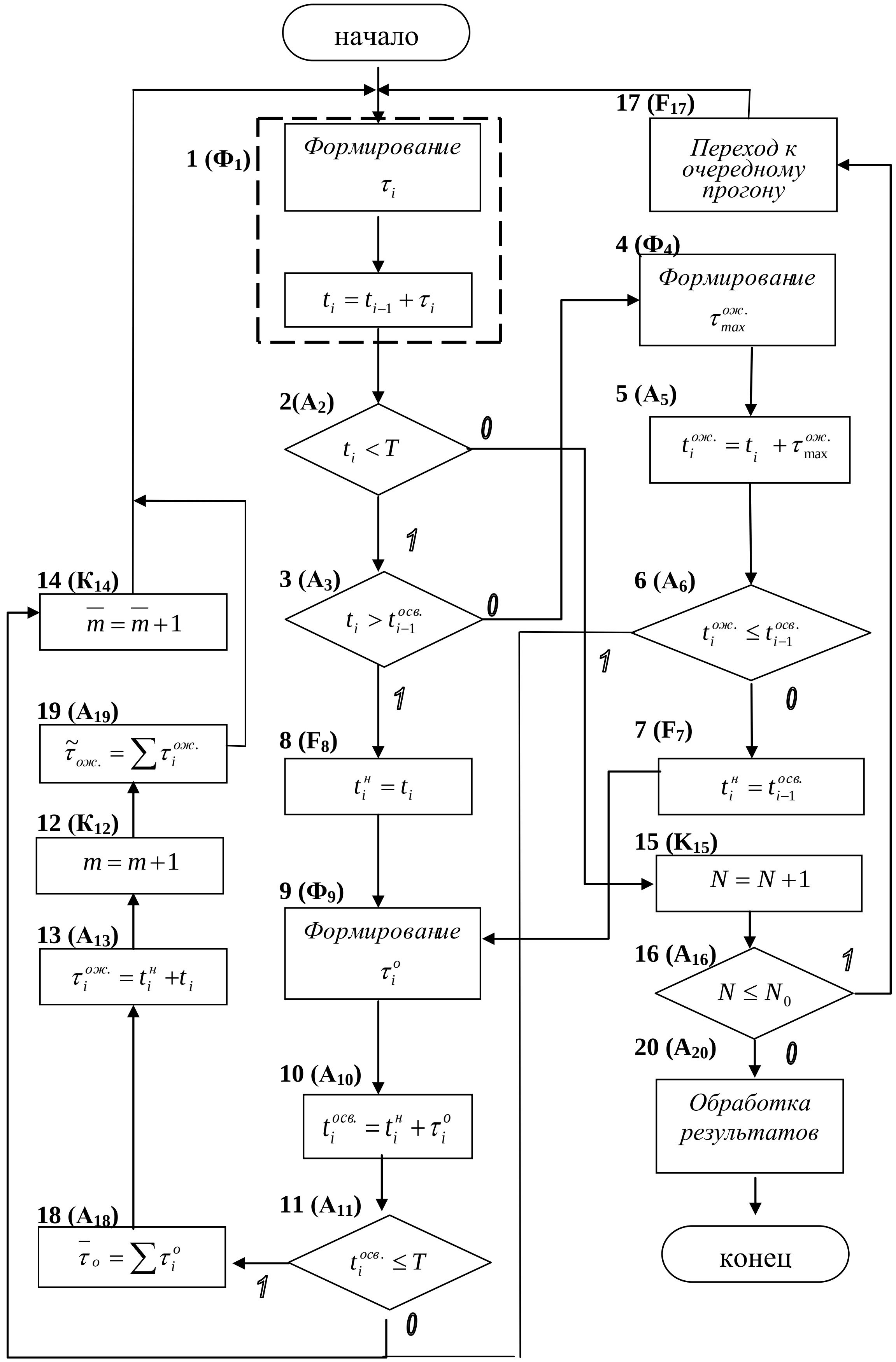
3. Имитаторы установки исходных данных.

Поскольку в системе необходимо провести несколько повторных прогонов, то в алгоритме предусматривается блок установки начальных условий при переходе к очередному прогону. Такой блок должен устанавливать: $t_i = 0$, $t_i^{осв.} = 0$, $t_i^н = 0$, $t^{;ож.} = 0$.

В итоге имитаторы задания исходных данных:

F_{17}: $t_i = 0$, $t_i^{осв.} = 0$, $t_i^н = 0$, $t^{;ож.} = 0$.

Структура алгоритма. Используя принцип особых состояний и последовательной проводки, составим алгоритм вида



Описание алгоритма. Оператором ***1*** формируется поток клиентов к оператору банка. Процедура проверки окончания модельного времени выполняется оператором ***2***. Если клиент обратился до истечения времени Т, он принимается оператором банка. Далее управление передаётся оператору ***3***, с помощью которого выясняется, свободен ли оператор банка. Если оператор банка свободен, управление передаётся оператору ***8***, который присваивает моменту начала обслуживания значение момента поступления в систему. Затем операторами ***9*** и ***10*** имитируется обслуживание клиента, а оператором ***11*** проверяется - достаточно ли оставшегося времени до конца работы оператора, чтобы обслужить клиента. Если «да», то управление передаётся операторам ***18***, ***13***, ***12***, ***19*** для сбора статистических данных. Далее клиент считается обслуженным, управление передаётся на вход оператора ***1*** для формирования нового поступления.

Если условие, записанное в операторе ***2***, не выполняется, то управление передаётся оператору ***15*** для подсчёта числа прогонов (смен работы оператора банка). Далее оператором ***16*** проверяется - по всем ли прогонам модели собрана статистика. Если «нет», то путём передачи управления оператору ***17*** переходят к имитации следующей смены работы оператора банка. Переход осуществляется путём установки нулевых начальных значений переменным, циркулирующим в алгоритме. От оператора ***17*** управление передаётся оператору ***1*** для формирования первого клиента новой смены.

Если условие оператора ***3*** не выполняется (клиент поступил в занятую систему), путём передачи управления операторам ***4*** и ***5*** имитируется постановка клиента в очередь. Далее управление передаётся оператору ***6***, проверяющему дождётся ли клиент, находящийся в очереди, своего обслуживания. Если «да», то управление от ***7*** передаётся операторам ***9***, ***10*** для имитации обслуживания. Если «нет», управление передаётся оператору ***14***, который ведёт подсчёт числа не обслуженных клиентов.

Если условие оператора ***11*** не выполняется, то управление передаётся оператору ***14*** для подсчёта числа не обслуженных клиентов.

Если условие оператора ***16*** не выполнено (все смены просмотрены), управление передаётся оператору ***20*** для обработки результатов и для вычисления искомых характеристик.

2.2.2.3. Проверка адекватности результатов имитационного моделирования

После построения модели ее необходимо проверить на соответствие получаемых данных данным объекта (адекватность модели).

Результаты модели считаются адекватными, если они близки к результатам, полученным на объекте при тех же условиях. Адекватность обычно проверяется по каждому из параметров **р** в отдельности, т.е. собирают статистические данные по одному и тому же параметру на объекте и на модели. Остальные параметры принимают не изменными.

Адекватность проверяется или по соответствию средних значений случайной величины (параметра), или по соответствию дисперсий случайной величины (параметра).

Проверка адекватности по средним значениям параметра. Для этого на объекте и на модели проводится N независимых испытаний. По результатам испытаний находят средние значения откликов (результатов) по каждому параметру р на объекте - $\overline{Y}^*_p$ и на модели - $\overline{Y}_p$ по формулам:

$$\overline{Y}_p = \frac{1}{N}\sum_{k=1}^{N} Y_{pk} \quad , \quad \overline{Y}^*_p = \frac{1}{N}\sum_{k=1}^{N} Y^*_{pk} \, .$$

Определяют оценку дисперсии на объекте - D^*_p и на модели - D_p по формулам: $D_p = \frac{1}{N-1}\sum_{k=1}^{N} \left(Y_{pk} - \overline{Y}_p\right)^2, \quad D^*_p = \frac{1}{N-1}\sum_{k=1}^{N} \left(Y^*_{pk} - \overline{Y^*_p}\right)^2.$

Вычисляют величину отклонения средних откликов на модели и объекте $\varepsilon = \left(\overline{Y}_p - \overline{Y}^*_p\right)$, определяют дисперсию полученной ошибки ε по формуле $D_{\varepsilon p} = \frac{D_p + D^*_p}{2}$, $t_{теор.}$ для выбранного уровня доверия α и V = 2(N-1) и $t_ф$ по формуле $t_ф = \frac{\varepsilon\sqrt{N}}{\sqrt{2D_{\varepsilon p}}}$. Если окажется, что $t_ф \le t_{теор.}$, то модель по параметру р адекватна. Если условие не выполняется, модель по данному параметру р не адекватна.

Проверка адекватности по дисперсии параметра. Для этого по каждому параметру р вычисляется дисперсия по выше изложенной формуле на объекте - D^*_p и на модели по формуле

$$D_p = \frac{1}{N-1}\sum_{k=1}^{N} \left(Y_{pk} - \overline{Y^*_p}\right)^2.$$

Адекватность модели проверяется на случайной величине F, которая подчиняется распределению Фишера. Она задана таблично и несёт информацию о соотношении дисперсии откликов модели и объекта.

Из табличных данных при выбранном уровне доверия α и степеням свободы $V_1=V_2=N$ определяется теоретическое значение $F_{теор.}$ Затем вычисляется фактическое значение $F_{факт.}=\frac{D_p}{D^*_p}$. Если $F_ф \le F_{теор}$, модель по дисперсии параметра р адекватна. Если условие не выполняется, модель по параметру р не адекватна. Если модель не адекватна по какому-либо параметру, переходят к ее калибровке.

2.2.2.4. Оценка устойчивости модели

При оценке адекватности модели как существующей, так и проектируемой системе реально может быть использовано лишь ограниченное подмножество всех возможных значений входных параметров (рабочей нагрузки и внешней среды). В связи с этим для обоснования достоверности получаемых результатов моделирования большое значение имеет проверка устойчивости модели. В теории моделирования это понятие трактуется следующим образом.

Устойчивость модели — это ее способность сохранять адекватность при исследовании эффективности системы на всем возможном диапазоне рабочей нагрузки, а также при внесении изменений в конфигурацию системы.

Устойчивость результатов моделирования можно рассматривать как признак, подлежащий оценке. Для проверки гипотезы об устойчивости результатов может быть использован критерий Уилкоксона.

Критерий Уилкоксона служит для проверки того, относятся ли две выборки к одной и той же генеральной совокупности (т. е. обладают ли они одним и тем же статистическим признаком). Например, в двух партиях некоторой продукции измеряется определенный признак, и требуется проверить гипотезу о том, что этот признак имеет в обеих партиях одинаковое распределение; другими словами, необходимо убедиться, что технологический процесс от партии к партии изменяется несущественно.

При статистической оценке устойчивости модели соответствующая гипотеза H может быть сформулирована следующим образом: при изменении входной (рабочей) нагрузки или структуры ИМ закон распределения результатов моделирования остается неизменным.

Проверку указанной гипотезы H проводят при следующих исходных данных: есть две выборки $X = (x_1,\ldots, x_n)$ и $Y = (y_1, ..., y_m)$, полученные для различных значений рабочей нагрузки; относительно законов распределения X и У никаких предположений не делается.

Значения обеих выборок упорядочиваются вместе по возрастанию. Затем анализируется взаимное расположение x_i и y_j. В случае $y_j < x_i$ говорят, что пара значений (x_i, y_j) образует инверсию.

Например, пусть для n = m = 3 после упорядочивания получилась такая последовательность значений: y_1, x_1, y_3, x_2, y_2, x_3; тогда имеем инверсии: (x_1,y_1), (x_2,y_1), (x_2,y_3), (x_3,y_1), (x_3,y_2), (x_3,y_3).

Подсчитывают полное число инверсий U. Если гипотеза верна, то U не должно сильно отклоняться от своего математического ожидания M:

$$M = \frac{n \cdot m}{2}\ .$$

От гипотезы отказываются, если |U-M| > Uкр (Uкр определяют по таблице для заданного уровня значимости).

2.2.2.5. Оценка чувствительности модели

. Очевидно, что устойчивость является положительным свойством модели. Однако если изменение входных воздействий или параметров модели (в некотором заданном диапазоне) не отражается на значениях выходных параметров, то польза от такой модели невелика (ее можно назвать «бесчувственной»). В связи с этим возникает задача оценивания чувствительности модели к изменению параметров рабочей нагрузки и внутренних параметров самой системы.

Такую оценку проводят по каждому параметру P_k в отдельности. Основана она на том, что обычно диапазон возможных изменений параметра известен. Одна из наиболее простых и распространенных процедур оценивания состоит в следующем.

1) вычисляется величина относительного среднего приращения параметра P_k:

$$\Delta P_k = \frac{(P_{k\max} - P_{k\min}) \cdot 2}{(P_{k\max} + P_{k\min})} \cdot 100\% \quad ;$$

2) проводится пара модельных экспериментов при значениях $P_k = P_{kmax}$ и $P_k = P_{kmin}$ и средних фиксированных значениях остальных параметров. Определяются значения отклика модели $Y_1 = f(P_{kmax})$ и $Y_2 = f(P_{kmin})$;

3) вычисляется ее относительное приращение наблюдаемой переменной Y:

$$\Delta Y = \frac{(Y_1 - Y_2) \cdot 2}{(Y_1 + Y_2)} \cdot 100\% \ .$$

В результате для k-го параметра модели имеют пару значений ($\Delta P_k, \Delta Y_k$), характеризующую чувствительность модели по этому параметру.

Аналогично формируются пары для остальных параметров модели, которые образуют множество $\{\Delta P_k, \Delta Y_k\}$.

Данные, полученные при оценке чувствительности модели, могут быть использованы, в частности, при планировании экспериментов: большее внимание должно уделяться тем параметрам, по которым модель является более чувствительной.

2.2.2.6. Калибровка модели

Калибровкой называется изменение структуры и состава модели таким образом, чтобы отредактированная модель бала адекватна по каждому из параметров ***p***, установленных в модели.

Калибровка проводится в 3 этапа. На каждом из этапов используются специальные средства оценки калибровки модели по каждому из параметров. Калибровка ведется в следующей последовательности. Сначала калибровку модели проводят путем сравнения распределения вероятности результатов, полученных на объекте -Y^*_{pk} и на модели - Y_{pk}. Если распределения совпадают, а адекватности не достигнуто, переходят к балансировке модели. На этом этапе значение параметров модели изменяют таким образом, чтобы результаты совпадали. Если балансировка не приводит к адекватности модели, переходят к оптимизации. Если оптимизация не приводит к адекватности, переходят к построению новой модели.

Калибровка сравнением законов распределения вероятности. Адекватность может быть не обеспечена в модели вследствие того, что законы распределения, принятые в модели, не соответствуют законам, которые действуют на объекте. Калибровка модели по этому пункту ведётся в следующей последовательности. В модель закладывают законы распределения, по предположению, близкие к объекту или же изменяют модель таким образом, чтобы законы совпадали. Для того чтобы проверить насколько законы совпадают, используется критерий согласия χ^2_p (критерий согласия Пирсона).

Для проверки по χ^2_p степени совпадения (согласия) распределений проводятся следующие вычисления. Теоретическое $\chi^2_{p теор.}$ определяется по таблице распределения при заданных значениях уровня доверия α и степенях свободы V = m-2, где m – число групп ряда распределения, составляющих статистическую совокупность (группа близких значений). Фактическое $\chi^2_{pф}$ определяют по формуле $\chi^2_{pф} = \sum_{i=1}^{m} \frac{(n_i^* - n_i)^2}{n_i^*}$, где n_i и n_i^* - частота попадания значений исследуемого параметра в i-ю группу соответственно в выбранном законе распределения и в теоретическом законе распределения (на объекте). Затем $\chi^2_{pф}$ сравнивается с $\chi^2_{p теор.}$. Если окажется, что $\chi^2_{pф} \le \chi^2_{p теор.}$, то выбранные законы распределения в модели совпадают с законами на объекте. Если законы не совпадают, то в модели задаются другие совпадающие законы.

Далее при совпадающих законах распределения проверяется адекватность модели. Если модель не адекватна, переходят ко 2-ому этапу – балансировке.

Калибровка балансировкой модели. Адекватность модели может быть не обеспечена из-за неправильно выбранной области значений параметров. Поэтому на этом этапе проводится поиск такой области значения по каждому из параметров, чтобы результаты совпадали. Выполняется это следующим образом. В модели и на объекте изменяется значение параметра p в широком диапазоне. При каждом значении параметра вычисляется сумма квадратов отклонения $S_p = \frac{1}{N} \sum_{k=1}^{N} (y^*_{pk} - y_{pk})^2$. Если найденная сумма квадратов S_p приняла минимальное допустимое значение, значит по параметру p осуществлена балансировка. Если баланс по какому-то параметру отсутствует, то путем изменения модели его добиваются. Далее проверяется адекватность сбалансированной модели. Если балансировка не даёт положительных результатов, переходят к оптимизации.

Калибровка оптимизацией модели. На этом этапе на основе корелляционно-регрессионного анализа (см. подпункт 5 пункта 6 раздела «Обработка результатов эксперимента») определяется параметр, который может искажать наибольшим образом структуру и состав модели (иметь сильную зависимость с большими значениями), например, p_i . Для его исключения модель перестраивается так, чтобы указанный параметр был незначительным, добиваясь при этом чтобы оценка значимости коэффициента a_1 в линейной регрессии $S_{pi} = a_0 + a_1p_i$, описывающей выходную функциональную зависимость модели, была минимальной. Если в результате изменения модели коэффициент a_1 становится незначимым, проводится проверка адекватности. Если модель и в этом случае окажется не адекватной, переходят к полной перестройке модели.

2.2.3. Язык моделирования GPSS World

При создании программ имитационного моделирования возникают задачи общие для широкого класса моделей. Это организация псевдопараллельного выполнения алгоритмов, динамическое распределение памяти, операции с модельным временем, имитация случайных процессов, введение массива событий, сбор и обработка результатов моделирования.

Для облегчения решения этих задач созданы специальные проблемно-ориентированные средства (программные системы), которые называются языками моделирования. По структуре и правилам программирования языки моделирования подобны алгоритмическим языкам высокого уровня, но имеют более высокий уровень. В настоящее время известно более 500 языков моделирования.

Одним из наиболее распространенных современных языков моделирования является GPSS World (General Purpose Simulating System-общецелевая система моделирования). Система GPSS ориентирована на класс объектов, процесс функционирования которых можно представить в виде множества состояний и правил перехода из одного состояния в другое, определяемых в дискретной пространственно-временной области. Примерами таких объектов являются вычислительные системы, сети ЭВМ, системы передачи сообщений и т.д. В качестве формальных моделей таких объектов используют системы массового обслуживания, автоматы, стохастические сети и т.д.

Система GPSS World содержит боле 40 программных блоков и следующие функциональные объекты: транзакты, очереди, устройства, функции и т.д.

Каждый из функциональных блоков и объектов представляет собой набор программных средств, выполняющих те или иные проблемно-ориентированные функции. Функциональный объект имеет свои стандартные числовые атрибуты (СЧА).

Функциональные объекты

1. **Транзакты (ТА)** имитируют пользователей системы, заявки, требования, обращения к системе и т.д. Каждый транзакт имеет индивидуальный номер и номер блока, в котором он находится и набор стандартных числовых атрибутов:

PS - приоритет,

X j – j-й номер ТА,

M i – время прохождения i-го ТА участка модели,

P j – j-й параметр ТА.

ТА создаётся специальным блоком GENERATE, проводится по системе (алгоритму) и выводится из системы оператором TERMINATE. При этом о нем собираются статистические данные.

2. **Устройства** имитируют процесс обработки ТА и включают в себя следующие операторы (программные блоки):

- SEIZE – занять устройство;
- ADVANCE - оператор обработки;
- RELEASE – вывод ТА из устройства.

Устройство имеет следующие СЧА:

- FR j – коэффициент использования j-го устройства;
- FT j – среднее время обработки j-м устройством;
- F j   – состояние устройства с номером j (0- устройство свободно, 1-устройство занято).

3. **Очереди** имитируют постановку, нахождение и вывод ТА из очереди. Очереди составляются двумя блоками:

- QUEUE – захват (вход в очередь);
- DEPART – выход из очереди.

Очереди имеют следующие СЧА:

- Q j – длина j-й очереди;
- QM j – максимальная длина j-й очереди;
- QA j – средняя длина очереди $(\overline{\ }_j)$;
- QT j – среднее время ожидания ТА в очереди $(\overline{\ }^{ож}_{\sum j})$ с учётом транзитных ТА.

4. **Функции** устанавливают функциональную зависимость между СЧА.

Функции бывают: - непрерывные C;
                 - дискретные D.

Функция задаётся набором пар точек, определяющих координаты кривой.

Операторы GPSS

Формат операторов:

Метка…Операция…Операнды…Комментарии

Метка – это символьный адрес перехода. Метка обязательно начинается с буквы и помещается в первый столбец. В шестом столбце записывается операция, состоящая из названия программного блока. Далее через пробел записываются операнды. В качестве операндов фигурируют СЧА, численные значения функций, соотношения. Если операндов несколько, между ними ставится пробел или запятая. После операндов пишутся комментарии. Поле операндов от комментария отделяется точкой с запятой.

Описание некоторых операторов GPSS

1. GENERATE (генерировать).

Программный блок предназначен для создания ТА с соответствующим набором атрибутов. Программный блок имеет только выход.

Формат: GENERATE A, B, C, D

В поле A записывается математическое ожидание интервала следования транзактов ($\bar{\tau}$).

В поле B записывается величина отклонения случайной величины от математического ожидания ($\Delta\tau$). По умолчанию отклонение равно 0. Так заполняется поле B в случае равномерного закона распределения.

В поле C устанавливается время появления на выходе генератора 1-го транзакта (t_1).

В поле D записывается число транзактов, которое должен сформировать генератор.

В случае, если закон отличается от равномерного, генератор задается в следующем формате. Например, генератор с экспоненциальным распределением (λ=0,25) и с использованием генератора случайных чисел RN1 запишется в виде:

GENERATE (Exponential(1,0,(1/0.25))).

2. TRANSFER (передать).

Оператор предназначен для имитации правил и условий обслуживания. С помощью этого оператора ТА можно передавать в различные программные блоки модели.

Формат: TRANSFER A, B, C, D

В поле A ставится условие (режим) передачи.

В поле B указывается номер следующего блока по условию. В поле C указывается номер блока, в который должен перейти ТА, если блок, указанный в поле B, занят. В поле D записывается индекс, используемый в условии «ALL».

Описание условий в поле A.

1. По этому условию «Пробел» ТА передаётся в блок, указанный в поле B, (безусловный переход). Если блок занят, ТА остаётся в предыдущем блоке.
2. .N - по этому условию через N- десятичное число записывается вероятность, с которой ТА будет переходить в блок, указанный в поле C.
3. «ALL» - по этому условию ТА пытается войти в блок B, если он занят, то в B+D, если тот занят, то в B+2D и т.д. до C.
4. «BOTH» - ТА пытается войти в блок B, если он занят, то в блок C, если блок C занят, то остаётся в предыдущем блоке.

Пример: 1. TRAN , MM1

Безусловная передача блоку с меткой MM1.

2. TRAN .300,MM1,MM2

С вероятностью 0.7 ТА будут переданы блоку с меткой MM1 и с вероятностью 0.3 в блок MM2.

3. SEIZE (занять).

С помощью оператора SEIZE производится ввод ТА в устройство, имитирующее обслуживание.

Формат: SEIZE A

В поле А указывается номер или имя устройства.

4. ADVANCE (задержать).

Блок имитирует обслуживание путём выделения транзакту определённого времени нахождения в блоке.

Формат: ADVACE A, B

В поле А устанавливается математическое ожидание времени обслуживания. В поле В – разброс времени обслуживания (аналогично блоку GENERATE).

Пример: 1. ADVACE 7, 2

Транзакт будет оставаться в блоке от 5 до 9 единиц времени.

2. ADVACE (Exponential (1,0,70)).

Транзакт будет задержан на время, равное произведению значения функции FN1 на 70.

5. QUEUE (встать в очередь).

Оператор имитирует постановку ТА в очередь.

Формат: QUEUE A, B

В поле А – имя очереди или её номер.

В поле В – число одновременно вводимых транзактов в очередь. По умолчанию 1.

6. DEPART (покинуть очередь).

Оператор предназначен для вывода ТА из очереди.

Формат: DEPART A, B

А – имя очереди; В – число выводимых ТА (по умолчанию 1).

7. RELEASE (освободить устройство, реализовать).

Оператор предназначен для вывода из устройства обслуженного ТА.

Формат: RELEASE A

В поле А – имя или номер устройства.

8. TEST (сравнение двух СЧА).

Оператор используется для передачи управления (ТА) по результатам сравнения двух СЧА.

Формат: TEST R A, B, C.

В поле R ставятся условия сравнения:

E – « = »; NE – «≠»; L – « < »; LE – « ≤ »; G – « > »; GE – « ≥ ».

В полях А, В записываются СЧА, подлежащие сравнению.

Если условие поля R выполнимо, то ТА входит в блок TEST и далее по программе. Если условие не выполнено, ТА пытается войти в блок, указанный в поле С (метка перехода). Если блок С занят, то ТА не заходит в TEST, а ожидает выполнение условия.

9. MARK (отметить).

Формат: MARK A

Блок записывает в параметр, указанный в поле А, значение текущего времени. Если поле А свободно, то в место времени создания ТА устанавливается текущее время.

Пример: 1. MARK

Заменяет время входа транзакта в модуль на текущее значение.

2. MARK 3

Записывает текущее время в параметр 3.

10. TERMINARE (завершить).

Блок предназначен для уничтожения транзактов, прошедших через модель.

<u>Формат:</u> TERMINARE A

В поле А указывается число уничтоженных транзактов, а в дальнейшем число, вычитаемое из содержимого оператора START.

<u>Примечание</u>. Если в поле А «пробел», то ТА уничтожается, но при этом содержимое оператора START не уменьшается. Используется это в тех случаях, если в модели необходимо задать время моделирования Т. Выполняется это следующим образом. После оператора TERMINARE с «пробелом» устанавливается группа операторов:

GENERATE T

В поле Т записывается время моделирования (момент появления 1-го ТА).

TERMINARE 1

STAR 1

Оператор START устанавливается в пункте меню Command/Start с указанием:

– числа прогонов модели,

– условия печати результатов (по умолчанию печатается стандартный отчёт, если стоит NP, печать отменяется)

– число прогонов, через которые осуществляется промежуточная распечатка результатов.

Работает START следующим образом. После каждого прогона из содержимого счетчика вычитается число, указанное в поле А оператора TERM. При достижении счетчика нулевого значения моделирование заканчивается и если есть разрешение, печатается отчёт.

Пример: 1. START 400

Выполнить модель до 400 прогонов с печатью стандартного отчета.

2. START 100, ,10

Выполнить модель до 100 прогонов и вывод отчета через каждые 10 завершений.

Более подробное описание GPSS World смотрите в /7/.

Пример моделирования средствами GPSS системы обслуживания с отказом

Пусть система состоит из одного обслуживающего аппарата. Время обслуживания подчиняется показательному закону распределения, с математическим ожиданием времени обслуживания $\overline{\tau^o}$ =70 с. На вход системы поступают требования с интенсивностью λ = 0,01 с$^{-1}$ (закон показательный).

Необходимо за T = 10000 с определить коэффициент загрузки обслуживающего аппарата $\bar{k}$, вероятность обслуживания R.

Для реализации задачи моделирования системы средствами GPSS World (воспроизведение процесса функционирования обслуживающего устройства) необходимы имитаторы:

- входного потока;
- обслуживающего аппарата;
- правила обслуживания.

Задание вспомогательных функций (сервисных функций) состоитв задании времени моделирования операторами. Сбор и обработку результатов моделирования GPSS проводит автоматически.

Структура моделирующего алгоритма имеет вид рис. 8.

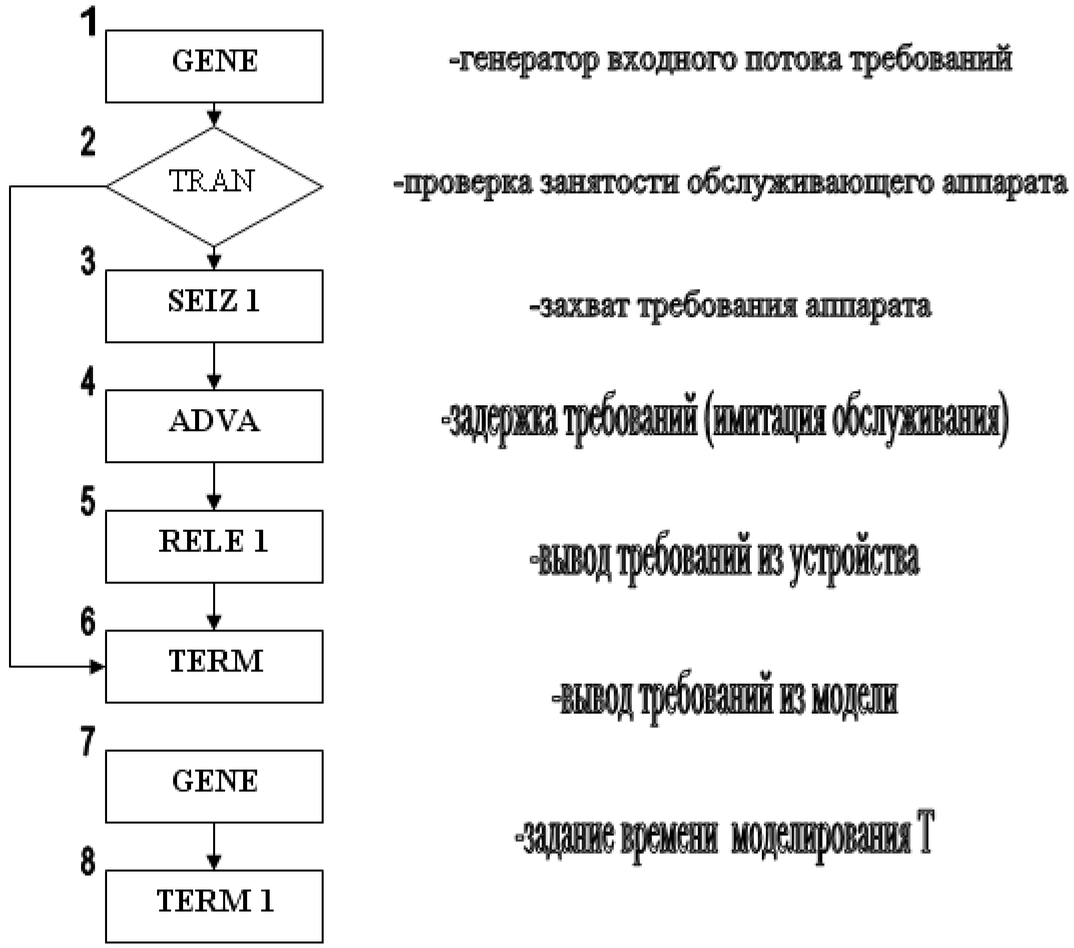


Рис. 8

1. Имитаторы входного потока:

GENERATE (Exponential (1, 0, 100)) - генератор требований с интервалом 100*FN$EXP.

2. Имитаторы обслуживания:

SEIZE 1- захват требований обслуживающим аппаратом 1;

ADVANCE (Exponential (1, 0, 70)) - обслуживание требований за время 70*FN$EXP;

RELEASE 1 - вывод требований из аппарата.

3. Имитаторы правил:

TRANSFER BOTH, MM2, MM1 – требования пытаются войти в блок с меткой MM2 (аппарат обслуживания), если он занят, то войти в блок с меткой MM1 (вывод из системы TERM).

4. Модельное время задается блоками:

GENERATE 10000

TERMINARE 1

Листинг программы модели:

```
        GENERATE (Exponential (1, 0, 100))
        TRANSFER BOTH, MM2, MM1
MM2     SEIZE  1
        ADVANCE (Exponential (1, 0, 70))
        RELEASE 1
MM1     TERMINARE
        GENERATE 10000
        TERMINARE 1
```

По окончании моделирования выводится стандартный отчёт, из которого находим коэффициент загрузки оператора $\bar{k}$=0,46. Так как число поступивших требований $N_{пост.}$=108 (созданных GENE), а обслуженных - $N_{обсл.}$=66 (поступивших в SEIZE1), то R = 66/108=0,61.

Литература

1. Советов Б.Я., Яковлев С.А. Моделирование систем. М.: Высш. шк., 2001.
2. Гультяев А.К. MatLab 5.2. Имитационное моделирование в среде Windows. Практическое пособие. – С.Пб.: Крона принт, 1999.
3. Морозов А.С. Моделирование технологических процессов и систем. Учеб. Пособие. Рязань, РГРТА, 2007.
4. Советов Б.Я., Яковлев С.А. Моделирование систем. Практикум. Учеб. пособие. М.: Высш. шк. 2003.
5. Язык моделирования GPSS World (студенческая версия).

Содержание

Состав курса:

группа	х37	х38	заочники
1. Лекции, ч	48	48	6
2. Лаб. раб., ч	16	16	6
3. К/раб.	7	7	контр/р.
4. Аттестация	экз.	экз.	экз.
5. Семестр	7	7	6

Адрес хранения: Кафедра/ Морозов/ Комплект по МС 2009/Лекции 2011