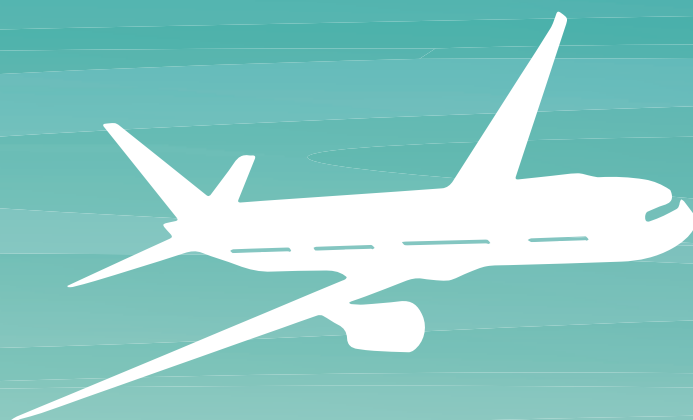




**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ГРАЖДАНСКОЙ АВИАЦИИ**

С.Н. Яблонский, С.В. Далецкий, К.Н. Матюхин

ИССЛЕДОВАНИЕ ОПЕРАЦИЙ И СИСТЕМНЫЙ АНАЛИЗ



**Москва
2017**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ВОЗДУШНОГО ТРАНСПОРТА

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ**

**«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ ГРАЖДАНСКОЙ АВИАЦИИ» (МГТУ ГА)**

**Кафедра технической эксплуатации летательных аппаратов
и авиадвигателей**

С.Н. Яблонский, С.В. Далецкий, К.Н. Матюхин

ИССЛЕДОВАНИЕ ОПЕРАЦИЙ И СИСТЕМНЫЙ АНАЛИЗ

Тексты лекций

Утверждено Редакционно-
издательским советом МГТУ ГА
в качестве учебного пособия

Москва-2017

УДК 629.735.083
ББК 517.8
Я14

Печатается по решению редакционно-издательского совета

Московского государственного технического университета ГА

Рецензенты: канд. техн. наук, доц. Е.Д. Герасимова (МГТУ ГА);
д-р техн. наук, проф. И.Г. Кирпичев (АК «РусДжет»)

Яблонский С.Н., Далецкий С.В., Матюхин К.Н.

Я14 Исследование операций и системный анализ: тексты лекций. — М.: МГТУ
ГА, 2017. — 80 с., 2 табл., 26 ил., лит.: 11 наим.

ISBN 979-5-903865-20-8

Данные тексты лекций содержат материалы учебно-методического характера, необходимые для освоения знаний и умений по предмету «Исследование операций и системный анализ». Содержит базовую часть образовательной программы Техническая эксплуатация летательных аппаратов и двигателей.

Тексты лекций издаются в соответствии с рабочей программой учебной дисциплины «Исследование операций и системный анализ» по Учебному плану направления 25.03.01 всех форм обучения.

Рассмотрено и одобрено на заседании кафедры 31.01.2017 г. и методического совета 02.02.2017 г.

ББК 517.8
Св. тем. план 2017 г.
поз. 7

ВВЕДЕНИЕ

Исторически, в развитии научного знания и его приложении к практической деятельности наблюдаются две противоположные тенденции.

1. Дифференциация - когда при увеличении знаний и появлении новых проблем из более общих наук выделяются частные науки.

2. Интеграция – когда более общие науки возникают в результате обобщения и развития тех или иных разделов смежных наук и их методов.

В сегодняшнем мире знаний одним из направлений перестройки в высшем профессиональном образовании является преодоление недостатков узкой специализации, усиление междисциплинарных связей, а также развитие системного мышления.

В учебный план многих вузов введены общие и специальные курсы, реализующие эту тенденцию. К числу таких дисциплин принадлежит и наш курс «Исследование операций и системный анализ» - типично междисциплинарный курс, обобщающий методологию исследования сложных систем. В различных технических вузах подобные дисциплины имеют свои названия: «Исследование операций» или «Системный анализ», или «Системный анализ и управление», «Теория систем и системный анализ» и т.п.

Цель курса лекций «Исследование операций и системный анализ», – познакомить студентов с основными методами исследования операций (ИО) и системного анализа (СА).

Из всего многообразия методов ИО и СА мы будем использовать методы, имеющие прикладное значение при решении задач, решаемых воздушным транспортом и его эксплуатационными предприятиями и связанных главным образом с технической эксплуатацией ЛА. Задача этих методов - помочь специалистам, ответственным за принятие решения, провести анализ ситуации и, в конечном счете, остановиться на том или ином оптимальном варианте действий. Это означает, что прошли те времена, когда различные управленческие решения находились методом проб и ошибок, так как такое управление приводит к большим потерям, связанные с ошибками.

Как известно, ИО – наука, занимающаяся количественным обоснованием решений во всех областях целенаправленной человеческой деятельности. В данном курсе лекций, из всего многообразия методов ИО, основное внимание уделено марковским и полумарковским случайным процессам, системам массового обслуживания, сетевому планированию и управлению (СПУ).

Основные направления СА, обеспечивающие специальные дисциплины, связаны с методами определения эффективности и показателей качества систем, вероятностно-статистическими методами анализа и методами исследования операций.

Математические методы СА и ИО имеют много общего. Можно даже сказать, что прикладные математические методы исследования операций являются одним из важнейших инструментов системного анализа.

Лекция № 1.1

1 раздел. Теория систем и системный анализ

Тема № 1.1 Основы системного анализа

- 1.1.1 Из истории исследования операций и системного анализа
- 1.1.2 Исследование операций как методологическая основа принятия решений
- 1.1.3 Системный анализ как последовательная реализация системного подхода
- 1.1.4 Методология системного подхода

1.1.1 Из истории исследования операций и системного анализа

Развитие научного знания и его приложение к практической деятельности в XVIII-XIX вв. привело к возрастающей дифференциации научных и прикладных направлений. Возникло много специальных дисциплин, которые часто используют сходные формальные методы, но настолько преломляют их с учетом потребностей конкретных приложений, что специалисты, работающие в разных прикладных областях (*так называемые «узкие специалисты»*), перестают понимать друг друга. В то же время в конце XIX в. стало резко нарастать число комплексных проектов и проблем, требующих участия специалистов различных областей знаний.

Усложнилось управление экономикой стран, отдельных предприятий и организаций. Появилась потребность в специалистах *«широкого профиля»*, обладающих знаниями не только в своей области, но и в смежных областях, использовать аналогии, формировать комплексные модели. Понятие системы, ранее употреблявшееся в обыденном смысле, превратилось в специальную общенаучную категорию, начали появляться обобщающие научные направления, которые исторически возникали параллельно на разной прикладной или теоретической основе и носили различные наименования.

Роль интеграции наук, организации взаимосвязей и взаимодействия между различными научными направлениями во все времена выполняла философия – наука наук, которая одновременно являлась и источником возникновения ряда научных направлений.

Междисциплинарные научные направления, возникшие между философией и специальными дисциплинами, можно расположить, как представлено в табл.1.1 [2;10].

Основные вехи.

1. В 30-е гг. XX в. из философии выделяется обобщающее направление, названное *теорией систем*. Основоположителем этого направления считается Л. фон Берталанфи, хотя и являлся биологом по основной профессии, но первый доклад о своей новой концепции сделал на философском семинаре. При этом использовал в качестве исходных понятий терминологию философии. Основным новым понятием, введенным Берталанфи, было понятие *открытой системы* [10].

Междисциплинарные научные направления

Ф И Л О С О Ф И Я	
МЕЖДИСЦИПЛИНАРНОЕ НАПРАВЛЕНИЕ	<i>Тектология</i>
	Теория систем
	Системология
	Системный анализ
	Системотехника
	<i>Информационный подход</i>
	<i>Концептуальное мета моделирование и проектирование</i>
	<i>Ситуационное моделирование</i>
	<i>Синергетический подход</i>
	<i>Имитационное моделирование</i>
	Кибернетика
	Исследование операций
СПЕЦИАЛЬНЫЕ ДИЦИПЛИНЫ	

В этом плане не нужно забывать и работы нашего соотечественника А.А. Богданова (*Малиновского*), предложившего в 20-е гг. всеобщую организационную науку - **тектологию**. Создание всеобщей организационной науки (*тектологии*) – одно из главных научных достижений А.А. Богданова. Но, исторически сложилось так, что учение А.А. Богданова так и не было принято его современниками.

2. В нашей стране вначале теорию систем и системный подход активно развивали философы - А.И Уёмов, Ю.А. Урманцев, В.С. Тюхтин, В.Н. Садовский.

Однако философская терминология не всегда легко преломлялась в практической деятельности. Поэтому потребности практики почти одновременно со становлением теории систем привели к возникновению направления, названного **исследование операций** [10].

3. Считается, что зарождение нового самостоятельного научного направления - исследование операций, называемого в Англии операционным исследованием, приходится на период с 1935 по 1938г.

Это время формирования основных положений этого научного направления.

Из-за дислокации в отделе оперативных действий (*Operations*) новое научное направление и получила своё оригинальное название по имени адреса, на который высылались почта для отдела: *Operations/Research*.

Необходимо отметить, что распространение идей ИО совпало с развитием методов математического программирования. А они имели преимущества. Поэтому многие работы по ИО (*особенно зарубежные*) базировались на методах математического программирования и отражали это в методиках ИО.

Наряду с направлением, ориентированным на модели математического программирования, развивалось и направление ИО, базирующееся на матема-

тической статистике. Это направление развивала Е.С. Вентцель. Она включила в число методов ИО и теорию игр, и математическую логику, графы и другие методы дискретной математики.

4. Кроме того, потребности практики, привели к тому, что в 60-е гг. XX в. широкое распространение получили термины **системотехника, системология**.

Возник ряд родственных направлений – *имитационное моделирование, ситуационное управление, информационный подход и др.*

Системотехника, или в другой терминологии **теория больших систем**, охватывает вопросы проектирования, создания, испытания и эксплуатации сложных систем. При разработке сложных систем возникают проблемы, относящиеся не только к свойствам их составных частей (*элементов, подсистем*), но также и к закономерностям функционирования объекта в целом (*общесистемные проблемы*). Появляется широкий круг специфических задач, таких, как определение общей структуры системы, организация взаимодействия между подсистемами и элементами, учёт влияния внешней среды, выбор оптимальных режимов функционирования, оптимальное управление системой и т. д.

5. Применительно к задачам управления в 50-е гг. XX в. более широкое распространение получил термин **кибернетика**, принятый Н. Винером для названия новой науки об управлении в живых организмах и машинах.

В нашей стране вначале кибернетика не признавалась. Ее называли лженаукой или потом мягче *«художественным приемом»*. А после признания кибернетики этот термин использовался в период становления работ по автоматизации управления как обобщающий для названия всех системных направлений. В связи с неоднозначной трактовкой термина, этот термин в настоящее время используется в более узком смысле – как одно из направлений теории систем, занимающееся процессами управления техническими объектами. А для обобщения научных направлений, связанных с исследованием и проектированием сложных систем в 70-х гг. XX века, используется термин **системные исследования**, а иногда сохраняется термин **системный подход** [10].

Надо иметь в виду, что «между разными направлениями системных исследований не существует принципиальных отличий. Вместе с тем каждое из этих направлений характеризуется определенными признаками и оттенками» [8].

1.1.2 Исследование операций как методологическая основа принятия решений

Мы будем использовать понятия «исследование операций» и «теория исследований операций» как синонимы.

Как известно, в терминах теории исследования операций искусственная система создается для выполнения операции. При этом функционирование системы представляется как некоторый процесс, который включает взаимосвязанные работы и действия, объединенные общим замыслом и единой целью. Именно такие целенаправленные процессы получили общее название **«опера-**

ция». Природа их может быть различной - это могут быть производственные процессы, коммерческие мероприятия, административные решения, военные действия, и т.д., и т.п., и пр., пр... .

Круг проблем, изучаемых ИО, весьма неоднозначен [4]. Иногда исследование операций понимают очень широко, включая в него ряд чисто математических методов, иногда, наоборот, очень узко – как практическую методику решения с помощью экономико-математических моделей строго определенного перечня задач.

Типичная ситуация такова: предполагается какое-то целенаправленное мероприятие (*система действий*), которое можно организовать тем или иным способом, т.е. выбрать какое-то «решение» из ряда возможных вариантов. Каждый вариант обладает какими-то преимуществами и какими-то недостатками, причем в силу сложности обстановки не сразу ясно, какой из них лучше (*предпочтительней*) других и почему. Подобного рода задачи сплошь и рядом возникают в различных областях практики, сама жизнь толкает нас на необходимость их решения.

Подход к этим задачам с общих, а не с узковедомственных позиций имеет ряд преимуществ: он расширяет кругозор исследователя, обеспечивает взаимопроникновение и взаимообогащение научных методов, подходов и приемов, выработанных в разных областях практики.

Операционные исследования, прежде всего, предназначены для предварительного количественного обоснования принимаемых решений. Кроме обоснования самих решений ИО позволяет сравнивать возможные варианты (*альтернативы*) организации операции, оценить возможное влияние на результат отдельных факторов, выявить «узкие места». Таким образом, сущность задач исследования операций – поиск путей рационального использования имеющихся ресурсов для реализации поставленной цели.

Методы исследования операций, как и любые математические методы, всегда в той или иной мере упрощают, огрубляют задачу, отражая нелинейные процессы линейными моделями, стохастические системы – детерминированными и т.д. Поэтому не следует ни преувеличивать значение количественных методов ИО, ни преуменьшать его, ссылаясь на примеры неудачных решений.

Сегодня мы будем говорить, что исследование операций - это направление, которое занимается количественным обоснованием решений во всех областях целенаправленной человеческой деятельности.

Несмотря на широкое распространение методологии ИО в различных прикладных областях все же исходная терминология этого направления (*в частности, само понятие «операция»*) часто трудно интерпретируется в условиях сложных технических комплексов, в экономических задачах и т.п. Поэтому наиболее конструктивным из направлений системных исследований в настоящее время считается системный анализ. Системный анализ – это прикладное направление теории систем. [2].

1.1.3 Системный анализ как последовательная реализация системного подхода

Необходимо учитывать, что и сегодня термин «системный анализ» трактуется в публикациях неоднозначно [10]. В одних источниках он определяется как «приложение системных концепций к функциям управления, связанным с планированием». В других - как синоним термина «анализ систем» (*Э. Квейд*) или термина «системные исследования» (*С. Янг*). Однако независимо от того, применяется он только к определению структуры целей системы, к планированию или к исследованию системы в целом, включая и функциональную и обеспечивающую части, работы по системному анализу существенно отличаются тем, что в них всегда предлагается методология проведения исследования, делается попытка выделить этапы исследования и предложить методику выполнения этих этапов в конкретных условиях. В этих работах всегда уделяется особое внимание определению целей системы, вопросам формализации представления целей. Некоторые авторы даже подчеркивают это в определении: системный анализ - это методология исследования целенаправленных систем (*Д. Киланд, В. Кинг*).

Термин «системный анализ» впервые появился в 1948г. в работах корпорации RAND в связи с задачами военного управления, для решения которых была разработана методика ПАТТЕРН, основанная на построении и анализе «дерева целей». В отечественной литературе термин получил распространение после перевода книги С. Оптнера «Системный анализ деловых и промышленных проблем» в 1969г. Введен в учебные планы вузов в конце 1980-х гг. Ф.И. Перегудовым (*зам. Министра высшего и среднего образования СССР того периода*).

Сегодня мы говорим: системный анализ представляет собой новое научное направление интеграционного типа, которое разрабатывает системную методологию решения сложных прикладных проблем и занимает определенное место в структуре современных системных исследований. Опирается на системный подход, а также на ряд математических дисциплин и современных методов управления. Термин «системный анализ» иногда употребляется как синоним системного подхода.

Согласно классификации, предложенной Г. Саймоном, все множество проблем в зависимости от глубины их познания подразделяется на 3 класса:

1. хорошо структурированные или количественно выраженные проблемы, которые поддаются математической формализации и решаются с использованием формальных методов;
2. неструктурированные или качественно выраженные проблемы, которые описываются лишь на содержательном уровне и решаются с использованием неформальных процедур;
3. слабоструктурированные (*смешанные проблемы*), которые содержат количественные и качественные проблемы, причем качественные, малоизвестные и неопределенные стороны проблем имеют тенденцию доминирования.

Для решения хорошо структурированных количественно выраженных проблем широко используется известная методология ИО, которая состоит в построении адекватной математической модели и применении методов отыскания оптимальной стратегии управления целенаправленными действиями.

В операционном исследовании выделяются основные этапы:

1. Определение конкурирующих стратегий достижения цели.
2. Построение математической модели операции.
3. Оценка эффективностей конкурирующих стратегий.
4. Выбор оптимальной стратегии достижения целей.

Появление системного анализа в методологии обоснования управленческих решений знаменует переход от решения хорошо структурированных, формализуемых проблем, к решению проблем слабоструктурированных, возникающих в условиях неопределенности и содержащих неформализуемые элементы, не переводимые на язык математики [8].

Академик Н.Н. Моисеев приводит, по его выражению, довольно узкое определение системного анализа: «Системный анализ - это совокупность методов, основанных на использовании ЭВМ и ориентированных на исследование сложных систем - технических, экономических, экологических и т.д. Результатом системных исследований является, как правило, выбор вполне определенной альтернативы: плана развития региона, параметров конструкции и т.д. Поэтому истоки системного анализа, его методические концепции лежат в тех дисциплинах, которые занимаются проблемами принятия решений: исследование операций и общая теория управления».

1.1.4 Методология системного подхода

Сегодня важно знать и понимать, что одной из наиболее существенных тенденций последних десятилетий в развитии и совершенствовании организации и управления, является становление так называемого системного подхода, а также широкое применение основополагающих идей и выводов теории систем.

В настоящее время вопрос ставится следующим образом: если мы хотим управлять профессионально, а не «по наитию», то должны овладеть методологией системного подхода, понятийным аппаратом теории систем, а для этого и при помощи этого, прежде всего, радикально изменить свое профессиональное, управленческое мышление – сделать его системным.

Системный подход считается детищем XX столетия. Системные теоретики утверждают, что ***системный подход*** – это не есть набор каких-то руководств или принципов для управленцев, это – способ мышления по отношению к организации и управлению.

Сегодня мы говорим, что при всех разногласиях системные теоретики единодушны, по крайней мере, в одном. Системный подход – есть методология познания частей на основании ***целого*** и ***целостности*** в отличие от классического подхода, который ориентирован на познание целого ***через части***.

Здесь речь идет о том, что система обладает собственными свойствами не присущими ни одному из ее составляющих элементов. И что крайне важно - свойства системы не сводятся к простой сумме свойств ее элементов, а, напротив, сами элементы определяются *целым и лишь в его рамках* получают свое функциональное объяснение и оправдание, как бы «право на существование».

Иногда говорят об *эмерджентности* системы. Речь идет о том же, что при объединении частей в целое возникает нечто качественно новое, которое проявляется в виде интегрированного качества. Это новое качество существует, пока существует целое, а само явление возникновения нового качества называют эмерджентностью (*от англ. – emergence – появление нового*).

Иногда говорят о системном или **синергетическом** эффекте (*синергия – от греч. – сотрудничество, содружество*) организации (*коллектива*). Это когда все сотрудники мотивированы единой целью, сплочены, вдохновлены, то возникает совершенно иная степень организованности данной организации.

Таким образом, можно сформулировать: в основе системного подхода лежит рассмотрение объектов, процессов и явлений как систем.

Системный подход:

1) рассматривает любую организацию (*а авиакомпания – это организация*) как целостную совокупность взаимосвязанных подсистем, объединенных общей целью и находящихся во взаимосвязи с внешней средой;

2) предполагает учет влияния всех ключевых факторов, воздействующих на организацию, и акцентирует внимание на взаимосвязях между ее подсистемами.

Системный подход – это, прежде всего правильная организация мышления, заключающаяся в умении воспринимать окружающий мир и его проблемы не через узко избирательный фильтр сиюминутных выгод и устремлений, а через многогранную призму всесторонней оценки последствий решений для всех, кого они затрагивают, позволяющую видеть проблему в целом во всей ее сложности и полноте.

Системный подход в управлении позволяет избежать ситуаций, когда хорошее решение для одной части организации превращается в проблему для другой. Системный подход требует рассматривать проблему не изолированно, а в единстве связей с окружающей средой, постигать сущность каждой связи и отдельного элемента, проводить ассоциации между общими и частными целями. Все это формирует особый метод мышления, позволяющий гибко реагировать на изменения обстановки и принимать обоснованные решения.

Из всего многообразия методов ИО и СА мы будем использовать методы, имеющие прикладное значение при решении задач, решаемых воздушным транспортом и его эксплуатационными предприятиями и связанных главным образом с ТЭ ЛА. При этом будем иметь в виду, что наша дисциплина позволяет синтезировать разрозненно полученные ранее знания в целостное образование, ибо природа и любое дело не знают нашего разделения на науки.

Раздел 1. Теория систем и системный анализ

Тема № 1.2. Основы теории систем, основные понятия и определения

1.2.1 Понятие «система».

1.2.2 Классификация систем.

1.2.3 Структура системы. Виды структур.

1.2.4 Примеры систем воздушного транспорта и их краткая характеристика.

1.2.1 Понятие «система»

Термин система используют в тех случаях, когда хотят охарактеризовать исследуемый или проектируемый объект как нечто целое (*единое*), сложное, о котором невозможно сразу дать представление, показав его, изобразив графически или описав математическим выражением (*формулой, уравнением и.т.п.*) и желают подчеркнуть, что это что-то большое, сложное и при этом целое, единое [10].

Понятие системы подчеркивает упорядоченность, целостность, наличие определенных закономерностей. Центральным моментом общей теории систем является понятие «*система*». Термин «система» греческого происхождения и в буквальном переводе, как гласит БСЭ, означает – «*целое, составленное из частей*». Несмотря на то, что понятие «система» используется во всех областях науки и техники, до сих пор оно еще не имеет достаточно четкого определения. Специалисты насчитывают до 40 определений этого понятия.

Людвиг фон Берталанфи, считающийся основоположником теории систем определял систему как «комплекс взаимодействующих компонентов» или как «совокупность элементов, находящихся в определенных отношениях друг с другом и со средой» [10].

Некоторые системные теоретики утверждают, что понятию «система» вообще нельзя дать строгого определения. Его можно лишь описать и в дальнейшем принимать таковым на основе соглашения.

Определяющим принципом рассмотрения некоторой совокупности элементов как системы является появление у нее новых свойств, которых не имеют составляющие ее элементы. Этот принцип ранее мы определили – принцип эмерджентности.

В качестве **элемента** системы будем рассматривать объект, относительно самостоятельный и не подлежащий дальнейшему расчленению на данном уровне рассмотрения, выполняющий определенные функции, находящийся во взаимосвязи с другими объектами, составляющими систему. Элементами системы могут быть как подсистемы, так и ее компоненты. Таким образом, в общем случае можно сформулировать следующее определение:

Система – это совокупность элементов, находящихся в отношениях и связях друг с другом, которая образует определенную целостность и единство в процессе своего функционирования для достижения некоторой цели.

Целостность системы и ее «отделенность» от окружающего мира обеспечиваются тем, что взаимосвязи внутри системы существенно сильнее, чем связь какого-либо ее элемента с любым элементом, лежащим вне системы.

1.2.2 Классификация систем

Классификация - это разделение совокупности объектов на классы по некоторым наиболее существенным признакам.

Важно понять, что любая классификация всегда относительна. Сама классификация выступает в качестве инструмента системного анализа. Полной классификации систем в настоящее время нет, более того, не выработаны окончательно ее принципы. Один из возможных подходов к классификации систем представлен на рис.1.1 [1].

В [1] классификация систем представлена по признаку *объективности существования*. По этому признаку все системы можно разбить на две большие группы: материальные (*реальные*) и абстрактные (*символические*) системы.

Материальные системы состоят из естественных и искусственных объектов, т.е. изделий, оборудования, машин и т.п.

Абстрактные системы, по сути, являются моделями реальных объектов - это языки, системы счисления, идеи, планы, гипотезы и понятия, алгоритмы и компьютерные программы, математические модели.

Далее используется признак по происхождению. В зависимости от происхождения системы делятся на природные (*естественные*) и искусственные (*создаваемые*).

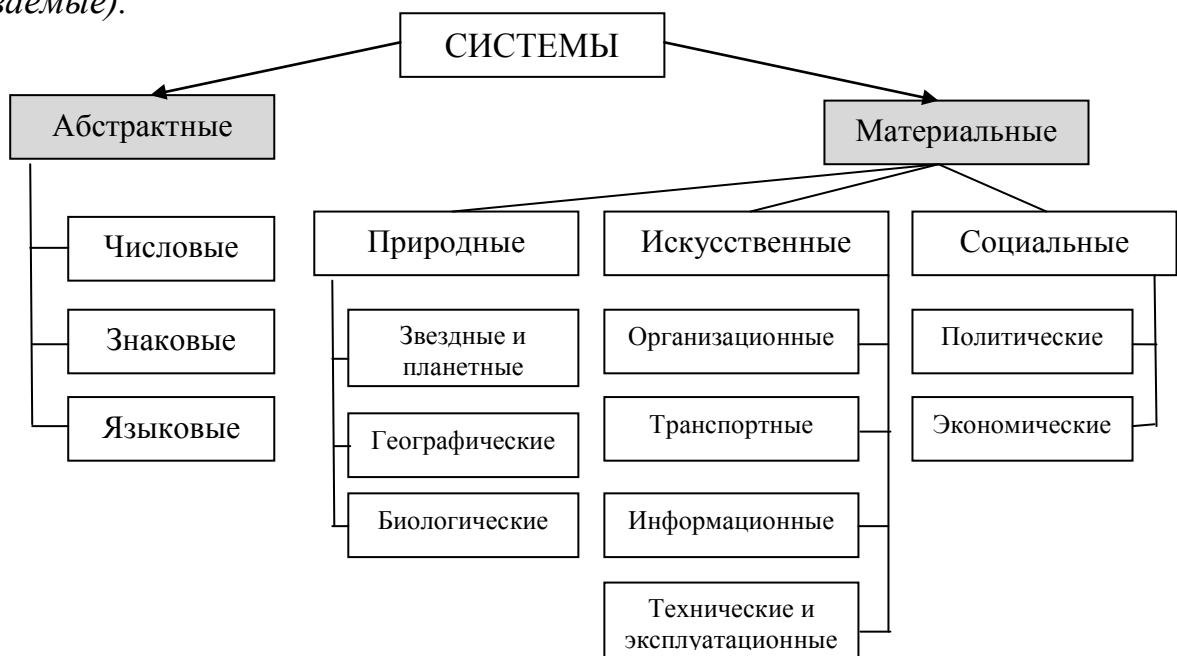


Рис.1.1 Классификация систем

Природные системы - это системы, объективно существующие в действительности. Эти системы возникли в природе без участия человека. Искусственные системы – системы, созданные человеком.

Диапазон этих систем весьма широк – от простейшего механизма до сложнейших комплексов; от отдельной лаборатории, воинского подразделения до министерства. Первые составляют подкласс технических систем, вторые – организационных.

Социальные системы – это тоже системы, создаваемые человеком и направленные на социальное развитие жизни людей. Это – политические, экономические, образовательные, медицинские, правовые и т.д.

Сегодня мы говорим, что научно-технический прогресс сопровождается непрерывным ростом сложности искусственных систем. Эти системы все в большей степени зависят от человека. Это привело к необходимости рассматривать человека в качестве одного из элементов сложных технических комплексов и, как следствие, к необходимости выделения особого подкласса искусственных подсистем – организационно-технических, являющихся объединением организационных и технических систем.

Особое место и интерес для нас, занимают сложные ОТС, которые относятся к классу искусственных систем. Иногда их называют по-разному: большие системы, системы «человек-машина», интегральные системы, системотехнические комплексы. Их особенностью является то, что помимо таких структурных элементов, как подразделения, коллективы людей, управленческий и технический персонал, ОТС включает в себя также сложные технические объекты, эксплуатация которых является основной целью функционирования системы.

Сложные ОТС характеризуются следующими наиболее важными признаками:

- 1) комплексным составом системы – наличием людей, машин (*технических подсистем*) и природной среды;
- 2) наличием единой цели функционирования;
- 3) наличием подсистем, каждая из которых имеет цель функционирования, подчиненную общей цели функционирования всей системы;
- 4) наличием иерархически связанных уровней управления;
- 5) наличием большого числа связей между подсистемами;
- 6) устойчивостью к воздействию внешних и внутренних возмущающих факторов.

Примером такой сложной организационно-технической системы в гражданской авиации является инженерно-авиационная служба (*ИАС*).

1.2.3 Структура системы. Виды структур.

Система может быть представлена простым перечислением элементов или «черным ящиком» (*моделью «вход-выход»*). Однако чаще всего такого представления недостаточно, так как требуется выяснить, что собой представляет объект, что в нем обеспечивает получение требуемых результатов, т.е., что обеспечивает достижение поставленной цели [10].

В этих случаях систему отображают путем расчленения на подсистемы и элементы с их взаимосвязями, которые могут носить различный характер, и вводят понятие структуры.

Структура (от латинского «structure» - строение, расположение, порядок) - это упорядоченная совокупность элементов и связей. Можно сказать, что структура характеризует сущность системы в ее статике, как некоторый временной срез.

Следует отметить, что, с точки зрения общей теории систем, структура – это только совокупность связей и отношений между элементами, сам набор элементов и их свойства рассматриваются отдельно. В рамках общей теории управления структура системы характеризуется как составом элементов с присущими им свойствами, так и совокупностью связей и отношений между ними.

Необходимо иметь в виду, что структуру того или иного объекта можно рассматривать:

- во-первых, как определенный тип сочетания его отдельных частей, и
- во-вторых, как упорядоченную совокупность связей между ними.

Однако часто она рассматривается по принципу «или – или», что является ошибочным подходом, ибо речь идет лишь о двух сторонах одной и той же медали, хотя и относительно самостоятельных.

Структура является важнейшей характеристикой системы, так как при одном и том же составе элементов, но при различном взаимодействии между ними меняется и назначение системы, и ее возможности. В подтверждение этого на рис.1.2 приведен пример пары различных систем, построенных из одних и тех же элементов.

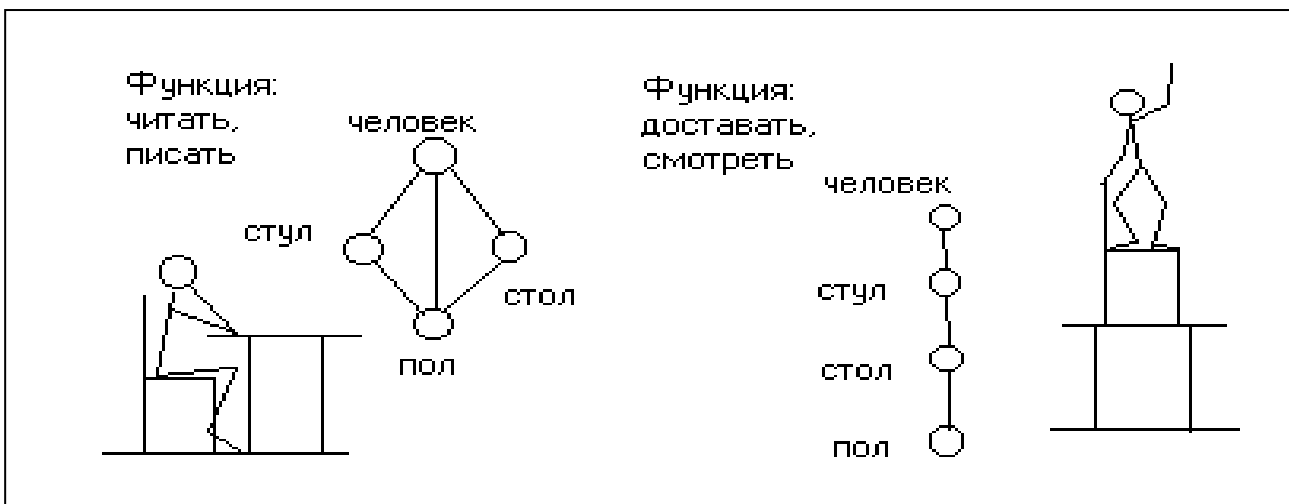


Рис.1.2 Пример пары различных систем

Рассмотрим типы структур, которые имеют место в ОТС. Это - функциональные, организационные, информационные.

Функциональная структура – структура, отражающая функциональные связи и взаимодействия между элементами. В функциональной структуре элементами являются функции, задачи и процедуры, а связи - информационные.

Организационная структура – упорядоченная совокупность взаимосвязанных частей организации – подразделений, обособившихся в процессе разделения труда. Структура организации отражает принятое в ней разделение работ между подразделениями, группами и людьми. Элементами организационной структуры являются коллективы людей и отдельные исполнители, а связи в этой структуре – информационные, соподчинения и взаимодействия.

Первые организационные структуры появились в армии еще до нашей эры.

Организационные структуры во многих организациях сложились больше исторически, чем в результате целенаправленных усилий по их формированию и улучшению. В управленческой литературе рассматриваются различные модели организационных структур и первой, систематически разработанной моделью, является бюрократическая структура организации, разработанная немецким социологом М. Вебером.

Концепция рациональной бюрократии, сформулированная М. Вебером, по крайней мере, в идеале, - это одна из наиболее полезных идей в истории человечества [9].

Основными принципами формирования «идеальной» бюрократической организации являются:

- 1) четкое разделение труда
- 2) иерархичность уровней управления
- 3) формализация и стандартизация деятельности

Формализация характеризует масштабы использования правил и процедур, которые определяют поведение работников в любой ситуации.

4) дух формальной обезличенности, с которым официальные лица выполняют свои должностные обязанности;

5) квалификационный отбор и выдвижение сотрудников по их деловым и профессиональным качествам.

Организационная структура, построенная в соответствии с этими принципами, получила название иерархической или бюрократической структуры.

Таким образом, бюрократическая организационная структура, характеризующаяся вышеуказанными характеристиками по Веберу, явно отличается от принятого понимания термина «*бюрократическая организация*», употребляемого нередко для обозначения больших и громоздких правительственных учреждений или частных предприятий, которые, как правило, в своей деятельности не учитывают индивидуальные запросы людей.

Вебер считал, что правила и инструкции исключают волю и желание отдельных людей. Жизнь показала, что добиться этого практически невозможно.

Если описание рациональной бюрократии покажется знакомым, то это объясняется тем, что большинство современных организаций представляют собой варианты бюрократии.

Для организационных структур в зависимости от характера управления могут быть виды структур: линейная структура; функциональная структура; линейно-функциональная структура: централизованная; децентрализованная.

Линейная структура – начальник осуществляет управление через линейных руководителей, т.е. имеется два уровня руководства. Иногда такую схему называют - сложная линейная структура или многоуровневая.

Функциональная структура - начальник руководит исполнителями через функциональных руководителей.

Линейно-функциональная структура.

Это традиционная структура. Такие схемы исторически возникли в рамках фабрично-заводского производства и явились соответствующей «организационной реакцией» на усложнившееся производство и необходимость взаимодействия с внешней средой. При линейно-функциональной структуре аппарат управления состоит из линейных и функциональных звеньев. Линейные звенья являются ведущими. Руководитель линейного подразделения осуществляет управление на основе единоначалия и несет персональную ответственность за деятельность подразделения.

При этом характерны две разновидности:

- централизованное управление - функциональные руководители дают свои распоряжения исполнителям только через линейного руководителя (*начальника*). В этой структуре функциональные руководители вступают в роли «советников».

- децентрализованное управление - функциональные руководители имеют право давать распоряжения своим исполнителям (*по своей специальности*) помимо линейных руководителей.

Информационная структура - структура, раскрывающая пути потоков информации, собирающих, обрабатывающих, передающих и использующих информацию. Информационная структура в качестве элементов включает формы существования и представления информации в системе, а связями являются операции преобразования информации в системе.

Для наглядности структуру системы часто изображают графически в виде элементов, взаимосвязи между которыми изображают соединительными линиями.

1.2.4 Примеры систем воздушного транспорта и их краткая характеристика

Гражданскую авиацию, как отрасль народного хозяйства, предназначенную для осуществления воздушных перевозок и другой летной работы, можно представить в виде **авиационной транспортной системы** (АТрС). Минимальной организационной структурной единицей ГА, сохраняющей все основные свойства и функции отрасли в целом, является эксплуатационное авиапредприятие.

АТрС представляет собой совокупность совместно действующих:

- летательного аппарата (ЛА);
- комплекса наземных средств по подготовке и обеспечению полетов;
- личного состава, занятого эксплуатацией и ремонтом ЛА и наземных средств;
- и системы управления процессом эксплуатации.

Она обладает всеми особенностями, присущими сложным техническим системам.

АТрС можно разделить на ряд функциональных самостоятельных систем:

- летной эксплуатации;
- технической эксплуатации;
- управления воздушным движением;
- коммерческой эксплуатации;
- аэродромной эксплуатации.

Особое место в системе технической эксплуатации занимает система ТО и Р. ТО и Р, как система (рис. 1.3), представляет собой совокупность взаимосвязанных звеньев - составных частей: объекта ТО и Р, производственно-технической базы, средств ТО и Р, инженерно-технического персонала, программы ТО и Р и эксплуатационно-технической документации (ЭТД).

Основным назначением системы ТОиР является поддержание летной годности ЛА. в процессе длительной эксплуатации, обеспечение исправности самолетного парка и условий для его эффективного использования по назначению.



Рис. 1.3 Структура системы ТО и Р ВС

Совершенство системы ТО и Р определяется прежде всего тем, насколько четко и полно обеспечивается взаимосвязь и взаимодействие между всеми составными частями системы, а также теснотой связи, которая с ее помощью обеспечивается между объективно существующими процессами повреждаемости конструкций (*разрушительными процессами*) при эксплуатации и процессами поддержания и восстановления их работоспособности. Чем теснее связь, тем совершеннее система ТО и Р.

Центральное место в системе ТО и Р занимает *программа ТО и Р* - основной документ, содержащий совокупность главных принципов и принятых разработчиком правил и решений по применению наиболее эффективных методов и режимов ТО и Р, реализованных в конструкции объектов при их проектировании и изготовлении и в эксплуатационной документации с учетом заданных требований и условий использования ЛА.

*Лекция № 1.3***Раздел 1. Теория систем системный анализ*****Тема № 1.3 Эффективность систем и показатели их качества***

1.3.1 Понятия «эффективность» и «качество» систем

1.3.2 Подходы к выбору показателей и критериев эффективности и качества.

1.3.1 Понятия «эффективность» и «качество» систем

Известно, что функционирование любой системы представляет собой некоторый *процесс*, который состоит в выполнении взаимосвязанных работ и действий, объединенных общим замыслом и единой целью.

Такие целенаправленные процессы получили общее название *«операция»*.

Природа их может быть различной - это могут быть производственные процессы, коммерческие мероприятия, административные решения, и т.д., и т.п.

Процесс функционирования любой системы можно представить, как множество операций, имеющих одну или несколько целей. Все операции должны быть упорядочены, взаимосвязаны и направлены на достижение общих целей системы. Каждую операцию можно разложить на множество более простых операций и элементарных действий. Вместе с тем любую операцию можно рассматривать как составную часть более широкой операции, проводимой вышестоящими руководителями.

В целом процесс функционирования системы также можно рассматривать как операцию, которая имеет определенную глобальную цель, состоящую из иерархии частных целей и задач.

Движущей силой любой осмысленной деятельности является наличие проблем, постоянно возникающих в жизни людей и организаций. Чтобы решить возникшую проблему, необходимо сформулировать одну или несколько целей и предпринять некоторые действия. Для достижения поставленных целей люди принимают решения. Принятие решений направлено не просто на достижение целей, а на выбор наилучших способов деятельности, которые хороши ровно настолько, насколько они эффективны.

«... в области разработки оценок качества функционирования систем значительные результаты получены в теории ИО, в терминах которой искусственная система создается для выполнения операции» [10].

Понятие *«операция»* включает, по меньшей мере, три основных компонента:

1. Управляющая деятельность руководителя, который организует и проводит операцию.

2. Активные средства (*люди, техника, материалы и другие ресурсы*), находящиеся в распоряжении руководителя.

3. Другие средства или системы, взаимодействующие с активными средствами в ходе проведения операции.

Указанные элементы раскрывают смысл понятия «операция» и отвечают на три вопроса: *как* действовать; *чем* действовать; *на что* воздействовать для достижения поставленной цели.

При этом существенное значение имеет понятие «цель операции».

Анализ определений «цели» показывает, что в понятие «цель» вкладывают различные оттенки – от идеальных устремлений, до конкретных целей – конечных результатов деятельности за определенный период.

Здесь важно понять, что любая операция - это процесс, в ходе которого происходит своеобразный «обмен» *ресурсов*, используемых для проведения операции, на некоторый *полезный эффект*.

В результате этого обмена система за приобретенную для себя пользу «расплачивается» определенным количеством ресурсов (финансовых, материальных и других) и временем, затраченным на проведение операции.

До сих пор не утихают споры ученых вокруг понятия «эффективность» и способов ее оценки. На сегодняшний день сложилось устойчивое представление о том, что -

Эффективность - это наиболее общее свойство любых целенаправленных процессов (*операций*), которое раскрывается через понятие «цель операции» и объективно выражается степенью достижения этой цели с учетом затрат ресурсов и времени.

Или по-другому, **эффективность операции** – это степень достижения цели.

Эффективность функционирования систем зависит от воздействия различных по своей природе факторов.

Принято выделять три основные группы факторов: качество активных средств; способы применения активных средств; условия применения активных средств операции.

Активные средства - это ресурсы, используемые для проведения операции.

В роли активных средств выступают люди, а также материальные, технологические, энергетические, информационные и финансовые ресурсы. Активные средства можно рассматривать как «инструменты», используемые для достижения цели операции.

В настоящее время в теории эффективности сложились два подхода к определению качества активных средств.

Качество как потенциальная эффективность операции.

Согласно этому подходу принято рассматривать типовую операцию, в которой некоторый объект (*например, техническая система*) или субъект является активным средством достижения поставленной цели.

После этого определяют так называемую потенциальную эффективность операции, которая достигается при *идеальном* способе применения данного объекта или субъекта.

Потенциальную эффективность операции иначе называют *эффективностью системы*, которая используется идеальным способом как активное средство операции.

Поскольку эффективность системы зависит не от способа применения, а лишь от свойств системы, то ее принимают за основную характеристику качества данной системы. Следовательно, эффективной следует считать систему, использование которой при идеальном способе применения позволяет достигнуть цели операции.

Таким образом, хотя понятие «эффективность» по определению относится не к системе, а к операции, оно может применяться для характеристики качества системы. Однако этот подход трудно использовать на практике, поскольку не всегда понятно, как оценить потенциальную эффективность той или иной операции. Поэтому для определения качества активных средств, как правило, используют более привычный подход на основе выявления их полезных свойств.

Качество как совокупность внутренних свойств системы

Этот подход состоит в понимании качества как набора *внутренних свойств* системы, характеризующих ее пригодность удовлетворять определенные потребности согласно своему назначению.

Другими словами, *качество* - это совокупность полезных свойств объекта или субъекта с точки зрения достижения целей конкретной операции.

При таком подходе об уровне качества системы судят по степени выраженности у нее того или иного свойства. Для измерения уровня качества используют *показатели качества*.

Способы применения активных средств - следующая важная группа факторов, влияющих на эффективность операций.

Эти факторы представляют собой порядок и приемы использования активных средств (*ресурсов*) для решения поставленных задач. Для каждого типа активных средств характерны свои, специфические способы применения. Более того, каждое активное средство имеет, как правило, множество альтернативных способов применения, из которых *следует выбрать наилучший или приемлемый*.

К ним относятся: способы распределения задач и полномочий; способы организации взаимодействия между участниками операции; варианты распределения дефицитных ресурсов; способы обмена информацией; способы контроля результатов операции и другие.

Условия применения активных средств характеризуются воздействием внешней и внутренней среды организации на ход и результат операции.

К ним относятся: природные факторы (*например, климатические условия, географическое положение, катастрофические явления природы*); активные действия других участников операции (*например, сознательное противодействие конкурентов, поведение потребителей*); ограничения (*ресурсные, экономические, юридические, экологические*).

Классификация факторов.

Все факторы, влияющие на эффективность операций, можно разделить на следующие группы: управляемые и неуправляемые; внешние и внутренние; определенные и неопределенные.

Управляемые факторы - это такие, на которые ЛПР может влиять по своему усмотрению при подготовке и проведении операции.

Как правило, управляемые факторы - это качество активных средств и способы их применения. Именно из множества значений управляемых факторов формируются управленческие решения руководителя.

Неуправляемые факторы – это такие, на которые ЛПР влиять не может и должен лишь учитывать их возможное влияние.

К ним, как правило, относятся условия применения активных средств. Например, действия конкурентов, законодательство, решения правительства, экономическая ситуация и другие.

По отношению к открытой системе факторы, влияющие на эффективность операций, могут быть *внешними* и *внутренними*.

Внешние факторы отражают влияние внешней среды, способствуя успешному решению задач (*полезные факторы*) или противодействуя ее целям (*вредные факторы*). **Внутренние факторы** представляют собой органическое сочетание таких составляющих внутренней среды, как цели, структура, внутриорганизационные процессы, технология, кадры, организационная культура. Это так называемые ситуационные факторы внутри системы.

Определенные факторы – это такие, значения которых известны с требуемой точностью. Влияние определенных факторов на результат операции можно предсказать достаточно точно.

Поэтому при наличии только определенных факторов в ситуации выбора каждой альтернативе можно сопоставить некоторый известный результат. Такие ситуации называют детерминированными и в этом случае говорят о принятии решений в условиях определенности.

К определенным факторам условно можно отнести качество активных средств, если оно известно, поведение партнеров и поставщиков, если с ними заключены соглашения, и некоторые другие.

Неопределенные факторы – это такие, значения которых не известны или известны не полностью. Такие ситуации называют *недетерминированными* и в этом случае говорят о принятии решений в условиях неопределенности. Например, в процессе управления организацией неопределенными факторами, как правило, являются действия конкурентов, потребительский спрос, отказы технических систем, поведение персонала организации и многие другие.

1.3.2 Подходы к выбору показателей и критериев эффективности и качества

Чтобы сравнивать между собой активные средства, способы и условия их применения, необходимо измерять или оценивать эффективность операций. Для этого используются **показатели эффективности**.

Чтобы выбирать наилучшие или подходящие активные средства, способы и условия их применения, необходимо четко знать цель проведения операции. Для этого используются **критерии эффективности**.

Рассмотрим эти понятия более подробно.

Для измерения уровня эффективности операций используются показатели эффективности.

Показатель эффективности есть мера степени достижения цели, т.е. соответствия реального и требуемого результатов операции. Формально показатель эффективности задается как некоторая функция, по значению которой судят об эффективности проводимой операции.

В общем виде для показателя эффективности можно выделить следующие факторы, от которых зависит этот показатель:

- неуправляемые факторы (*известные параметры системы и условия её функционирования* - a_1, a_2, \dots , которые являются постоянными и изменены быть не могут);

- неопределенные факторы (*неизвестные условия или факторы* - Y_1, Y_2, \dots);

- управляемые факторы (*элементы выбора в ходе решения задачи* - x_1, x_2, \dots).

Таким образом, в общем виде можно записать

$$W = W(a_1, a_2, \dots, Y_1, Y_2, \dots, x_1, x_2, \dots).$$

Показатели эффективности, которые используются в задачах принятия решений, имеют не только разный смысл, но также отличаются по своей *форме*.

Прежде всего, отметим, что показатели могут быть *единичными и комплексными*.

Единичный показатель количественно характеризует только одно свойство системы. **Комплексный показатель** отражает не менее двух свойств (*задач, целей*) системы.

В качестве примера единичных показателей можно назвать показатели безотказности технических систем – интенсивность отказов, вероятность безотказной работы, средняя наработка до отказа.

В каждой конкретной задаче принятия решения должен быть сформирован свой набор показателей эффективности, значения которых точно или приближенно оценивают альтернативные варианты действий. Подобные показатели позволяют руководству сопоставить реальные результаты деятельности с желаемыми и ответить на ряд важных вопросов.

Кроме того, на основе реальных или прогнозируемых значений показателей можно сравнить любую пару альтернатив и вынести одно из трех суждений: одна из альтернатив предпочтительнее, чем другая; обе альтернативы одинаковы по предпочтительности; альтернативы несравнимы между собой. Поэтому показатели эффективности играют роль *критериев сравнения* альтернатив в процессе принятия управленческих решений.

В связи с этим вместо термина «показатель» очень часто используют термин «критерий». Такая замена правомерна, если под критерием понимать именно критерий сравнения альтернатив, а не критерий эффективности операции.

Однако, чтобы не возникало путаницы, когда речь идет об измерении эффективности, лучше всегда использовать термин «показатель».

Сложнее дело обстоит с формированием комплексных (*обобщенных*) показателей. Для их формирования применяются различные способы. Если систе-

ма многоцелевая и нет оснований отдавать предпочтение какой-либо одной главной цели, то формируется совокупность показателей эффективности.

Например, для авиационной транспортной системы применяется следующая совокупность показателей: безопасность полетов; регулярность вылетов; интенсивность и экономичность использования АТ.

В этом случае формируется не комплексный показатель, а комплекс показателей.

Другим подходом является построение обобщенного показателя на основе аддитивных и мультипликативных преобразований частных показателей.

Аддитивная (от лат. *additives* – прибавляемый) и *мультипликативная* (от лат. *multiplico* – умножаю, увеличиваю) формы многомерных показателей обладают замечательным свойством – возможностью свести многомерные характеристики к одномерным показателям, особым образом, сформированным и отражающим многомерную информацию.

Аддитивные комплексные показатели формируются в виде суммы частных показателей

$$U = \sum \alpha_i w_i,$$

где α_i – положительные или отрицательные коэффициенты.

Положительные коэффициенты ставятся при тех частных показателях, которые желательно максимизировать, отрицательные – при тех, которые желательно минимизировать.

Оценку весовых коэффициентов каждого показателя можно осуществить, используя *методы экспертных оценок* – метод баллов или метод ранжирования (*ранговой корреляции*).

Общий вид *мультипликативного обобщенного показателя*:

$$U = \prod_{i=1}^k w_i^{\lambda_i},$$

где λ_i – некоторые вещественные числа (в частном случае $\lambda_i = 1$).

Отметим, что вид данного мультипликативного обобщенного показателя ничем не отличается от предыдущего по конструкции, так как можно перейти к

логарифмам, и тогда $\ln U = \sum_{i=1}^k \lambda_i \ln w_i$

Нередко в качестве обобщенного мультипликативного показателя берут показатель в виде дроби, в числителе которой ставят те показатели, которые желательно увеличить (*положительные показатели*), а в знаменателе – те, которые желательно уменьшить (*отрицательные показатели*).

$$U = \frac{W_1 W_2 \dots W_m}{W_{m+1} W_{m+2} \dots W_k}$$

Главным недостатком этого подхода является то, что, уменьшая знаменатель при незначительной величине числителя, можно обеспечить большое значение критерия, т.е. недостаток эффективности по одному показателю всегда можно скомпенсировать за счет другого.

Критерии эффективности.

Процесс выбора критерия эффективности функционирования (*операции*) является в значительной мере субъективным, творческим. Для принятия управленческих решений недостаточно знать, по каким показателям будет оцениваться эффективность операции. Чтобы принять решение, необходимо использовать определенные правила, позволяющие выбрать наилучший или приемлемый вариант действий из множества допустимых.

Таковыми правилами являются критерии эффективности. Критерий эффективности есть решающее правило, которое позволяет осуществить целенаправленный выбор из множества альтернатив. Критерий эффективности формулируется на основе выбранных показателей.

В отличие от показателя, который лишь оценивает количественно или качественно степень достижения цели, критерий эффективности позволяет вынести суждение о приемлемости решения. По этой причине критерии эффективности играют роль критериев выбора альтернатив в процессе принятия решений.

Критерий эффективности вводится на основе определенной концепции принятия решений человеком. Как известно, существуют две такие концепции - ***максимизации полезности*** и ***ограниченной рациональности***.

Каждой из них соответствует определенный *вид* критериев эффективности, который называется критерий оптимальности и критерий пригодности.

Критерий оптимальности.

Критерии оптимальности используются для принятия оптимальных решений в соответствии с концепцией максимизации полезности. Управленческое решение считается оптимальным, если оно обеспечивает максимальную эффективность операции.

Если для принятия решения используется скалярный показатель эффективности, то он обычно играет роль функции полезности, которую необходимо максимизировать. Если для принятия решения используется векторный показатель эффективности, то роль функции полезности обычно играет некоторое математическое выражение в виде комбинации частных показателей. Такое выражение иначе называют обобщенным показателем эффективности.

Если число частных показателей и число альтернатив велико, то при этом могут использоваться вспомогательные методы, облегчающие процесс поиска и выбор наилучшей альтернативы. В этом случае оптимальным считается наилучшее компромиссное решение, принятое человеком субъективно с учетом значений всех частных показателей эффективности. Таким образом, критерий оптимальности - это правило, согласно которому следует принимать только такие решения, которые обеспечивают максимальную эффективность операции.

Критерий пригодности.

Критерии пригодности используются для принятия удовлетворительных решений в соответствии с концепцией ограниченной рациональности. Управленческое решение считается удовлетворительным, если оно обеспечивает требуемую эффективность операции.

Раздел 1. Теория систем и системный анализ

Тема № 1.4 Экстремальные задачи системного анализа

- 1.4.1 Научный подход к принятию решения
- 1.4.2 Задачи выбора и задачи оптимизации
- 1.4.3 Однокритериальные и многокритериальные задачи оптимизации.
- 1.4.4 Множество Парето.

1.4.1 Научный подход к принятию решения

Итак, мы уже знаем, что по существу все системные направления (*а мы рассматривали: теория систем, системный подход, системология, системный анализ, системотехника, кибернетика, исследование операций*) ориентированы на исследование проблем в сложных системах (*в частности ОТС*). Результатом такого исследования является выбор вполне определенного решения.

Для всякого, кто хочет преуспеть в искусстве управления чрезвычайно важно понимание природы принятия решений.

Понятие «управленческое решение» имеет *два смысловых значения*.

В широком смысле под решением понимают процесс выбора одного или нескольких вариантов действий из множества возможных. Иначе говоря, в этом смысле термин «решение» означает *процесс* принятия решения. В **узком смысле** решение есть *результат конкретного выбора* варианта действий.

В любом случае справедлива следующая формула: **решение – это выбор альтернативы**. Принятия решений не удастся избежать никому. Даже отказ от выбора - это тоже выбор.

Одна из древних мудростей гласит: *«Единственная реальная проблема в жизни состоит в том, чтобы решить, что делать дальше»*.

Сегодня мы будем говорить, что принятие решений – часть каждодневной работы любого руководителя. При принятии решения следует учитывать два существенных момента [9]:

Первый момент заключается в том, что принимать решения, как правило, сравнительно легко. Гораздо труднее принять правильное решение. Классики утверждают, что *более чем что-либо другое* компетентность в данной области отличает руководителя от не руководителя и, что еще важнее, эффективно работающего руководителя от его не эффективно работающего коллеги.

Второй момент состоит в том, что принятие решения – *это психологический процесс*. Все мы по опыту знаем, что человеческое поведение не всегда логично. Иногда нами движет логика, иногда чувства. Поэтому неудивительно, что способы, используемые руководителем для принятия решений, варьируются от спонтанных до в высшей степени логичных. В силу этого процесс принятия решений может быть основан на различных подходах. В практике управ-

ленческой деятельности различают три подхода: *интуитивный подход, подход, основанный на суждениях и подход, имеющий рациональный характер* [9].

Чисто интуитивное решение – это выбор, сделанный только на основе ощущения того, что он правилен. ЛПР, не занимается сознательным взвешиванием «за» и «против» по каждой альтернативе и даже не нуждается в понимании ситуации, просто человек делает выбор. То, что мы называем озарением или шестым чувством, и есть интуитивные решения.

Решение, основанное на суждении, - это выбор, обусловленный общими знаниями или накопленным опытом. Опираясь на опыт и здравый смысл, ЛПР выбирает, как правило, альтернативу, которая принесла успех в прошлом. Таким образом, условия - *общие знания, опыт, здравый смысл*.

Такой подход, характеризуется тем, что качество решения основывается, прежде всего, на предыдущем опыте руководителя (ЛПР). Подход не дает хороших результатов, когда опыт ЛПР небольшой, а ситуация совершенно новая. А, как говорят, классики, здравый смысл встречается редко [9].

Если проблемная ситуация не так очевидна, достаточно сложна, а ее решение неоднозначно, то используется **подход, имеющий рациональный характер**. В данном случае термин «рациональный» характеризует, прежде всего, способ разработки решения, а не его качество.

Рациональные решения принимаются с помощью многоэтапного аналитического процесса на основе и по результатам объективного анализа, как самой проблемы, так и альтернатив решения. Рациональные решения считаются наиболее обоснованными, так как используют все доступные человеку механизмы - *интуицию, логику и расчет*. В этом случае говорят о применении так называемого **научного подхода к принятию решения**.

В науке управления научный подход подразумевает наличие определенной структуры процесса принятия решений и использование различных методов и моделей принятия решений. В настоящее время разработано достаточно много различных подходов к выделению этапов принятия решений.

В результате анализа различных подходов, напрашивается вывод, что предложенные в них этапы принятия решения различаются в деталях, но в целом дают верное описание логики развития этого процесса.

Стэнфорд Оптнер в известной книге «*Системный анализ для решения, деловых и промышленных проблем*» предложил 12 этапов принятия решений; Стэнли Янг в работе «*Системное управление организацией*» предлагает десять этапов принятия решений. Мы рассмотрим этапы принятия решений, взяв за основу подход, предложенный в [9].

М. Мескон, М. Альберт и Ф. Хедоури определяют следующие «этапы рационального решения проблем»:

- 1) диагностика проблемы (*уяснение задачи и оценка обстановки*);
- 2) формулировка ограничений и критериев принятия решения;
- 3) определение альтернатив (*выработка замысла решения*);
- 4) оценка альтернатив (*количественное обоснование вариантов замысла*);

5) выбор альтернативы (*выбор варианта решения*);

6) реализация и обратная связь (*формулировка решения и его документирование*).

1 этап. Определение (*диагностика*) проблемы.

Существуют два подхода к определению понятия «проблема».

Во-первых, проблема понимается как *расхождение между действительным и желаемым* при неизвестных способах преодоления этого расхождения. Будем называть их проблемами *функционирования* организации.

Во-вторых, проблема часто понимается как потенциальная возможность. В этом смысле проблема означает расхождение между действительным и потенциально возможным при неизвестных способах преодоления этого расхождения. Назовем их проблемами развития организации.

2 этап. Формулировка ограничений и критериев принятия решения.

Когда руководитель диагностирует проблему с целью принятия решения, он должен отдавать себе отчет в том, что именно можно с нею сделать. Многие возможные решения проблем организации не будут реалистичными, поскольку либо у руководителя, либо у организации недостаточно ресурсов для реализации принятых решений. Кроме того, причиной проблемы могут быть находящиеся вне организации силы - такие, как законы, которые руководитель не властен изменить.

3 этап. Определение альтернатив (*выработка замысла решения*)

В идеале желательно выявить все возможные действия, которые могли бы устранить причины проблемы и, тем самым, дать возможность организации достичь своих целей. Тем не менее, на практике руководитель редко располагает достаточными знаниями или временем, чтобы сформулировать и оценить каждую альтернативу.

4 этап. Оценка альтернатив (*количественное обоснование вариантов замысла*).

При оценке возможных решений руководитель пытается спрогнозировать то, что произойдет в будущем. Будущее всегда неопределенно. Множество факторов может помешать воплощению намеченного. Поэтому важным моментом в оценке является определение вероятности осуществления каждого возможного решения в соответствии с намерениями.

5 этап. Выбор альтернативы (*выбор варианта решения*).

Необходимо иметь в виду – оптимальное решение это - наилучшее. Оно выбирается на основе концепции максимизации полезности. Удовлетворительное решение (*рациональное*) – на основе концепции ограниченной рациональности.

6 этап. Реализация и обратная связь (*формулировка решения и его документирование*). Принятое и сформулированное решение обычно оформляется в виде документа (*письменного, графического*), а при дефиците времени может формулироваться устно.

Перечисленные выше этапы отражают *нормативную структуру* процесса принятия решений и отвечают на вопрос: как надо принимать решения в сложных, проблемных ситуациях? Надо иметь в виду, что в реальных ситуациях руководители очень редко последовательно и в полном объеме «проходят» через

все эти этапы. Данное явление хорошо известно в теории и практике управления как *феномен расхождения нормативного и реального процессов принятия решений*.

Некоторые этапы могут выполняться людьми автоматически и неосознанно или вообще «выпадать» из общего процесса принятия решения. С другой стороны, может просто отсутствовать сама необходимость в выполнении отдельных этапов. Это имеет место, если решение принимается в повторяющихся, т.е. достаточно стереотипных ситуациях. В подобных ситуациях наиболее эффективны *решения*, основанные на логических суждениях и аналогиях. Такие решения принимаются быстро, имеют достаточно высокое качество и сравнительно небольшой риск.

1.4.2 Задачи выбора и задачи оптимизации

Выбор как реализация цели. Выбор является действием, придающим всей деятельности целенаправленность. Именно выбор реализует подчиненность всей деятельности определенной цели. Способность сделать правильный выбор - очень ценное качество, которое присуще людям в разной степени.

Великие полководцы, выдающиеся политики, гениальные инженеры и ученые, талантливые администраторы отличались и отличаются от своих коллег или конкурентов, прежде всего умением принимать лучшие решения, делать лучший выбор. Задачи выбора чрезвычайно многообразны, различны и методы их решения. К настоящему моменту сложились три основных языка описания выбора. Самым простым и наиболее развитым (*и, быть может, поэтому чаще употребляемым*) является критериальный язык.

Критериальный язык описания выбора.

Такое название языка связано с основным предположением, состоящим в том, что каждую отдельно взятую альтернативу можно оценить конкретным числом (*значением критерия*), и сравнение альтернатив сводится к сравнению соответствующих им чисел.

Из факторов, определяющих показатель в общем виде выбору подлежат значения величин x_1, x_2, \dots , т.е. управляемые факторы. В общем виде для показателя эффективности *ранее* выделялись следующие факторы:

- неуправляемые факторы (*известные параметры системы и условия её функционирования - a_1, a_2, \dots , которые изменены быть не могут*);
- неопределенные факторы (*неизвестные условия или факторы - Y_1, Y_2, \dots*);
- управляемые факторы (*элементы выбора - x_1, x_2, \dots*).

Сам выбор осуществляется из некоторого множества альтернатив.

Сформулируем в общем виде постановку задачи выбора для случая, когда имеется один элемент выбора x_1 , и выбор необходимо сделать между двумя альтернативами x_1^1 и x_1^2 .

Пусть соответствующие целевые функции суть:

- $W_1^1(a_1, a_2, \dots, Y_1, Y_2, \dots, x_1^1)$;
- $W_1^2(a_1, a_2, \dots, Y_1, Y_2, \dots, x_1^2)$.

Мы говорим, что альтернатива x_1^1 предпочтительнее x_1^2 , если $W_1^1 > W_1^2$

Задачи оптимизации.

Явно или неявно мы встречаемся с оптимизацией в любой сфере человеческой деятельности от сугубо личного до самого высокого общегосударственного уровня. Оптимизация (*от лат. optimus-наилучший*).

Под оптимизацией обычно понимают целенаправленную деятельность, заключающуюся в получении наилучших результатов при соответствующих условиях. Употребляя слово «оптимизация», нужно обязательно указывать, по какому критерию она понимается. Постановка задачи оптимизации предполагает наличие ее объекта, набора переменных x_1, x_2, \dots, x_n , описывающих данную задачу, а также условий (*часто называемых ограничениями*), которые характеризуют приемлемые значения независимых переменных a_1, a_2, \dots .

«Главный принцип должен оставаться в силе: *что-то должно оптимизироваться, а что-то задаваться*. Соответственно, первое что-то называется критерием оптимальности, а второе – ограничениями. Без этих двух элементов задачи оптимизации не бывает, хоть ты лопни от учености!»

Еще одной обязательной компонентой описания задачи оптимизации служит скалярная мера «качества», носящая название целевой функции, и зависящая каким-либо образом от переменных оптимизации.

Решение оптимизационной задачи - это поиск определенного набора значений переменных, которому отвечает оптимальное значение целевой функции.

В общем случае это можно представить в следующем виде:

$$F(x) = f(x_1, x_2, \dots, x_n) \rightarrow \max (\min);$$

где - $F(x)$ – целевая функция (*функция полезности*), заданная на множестве альтернатив.

Наиболее просты, с математической точки зрения, случаи, когда целевая функция задается явной формулой и является при этом дифференцируемой функцией. В этом случае для поиска точек локального экстремума может быть использована производная.

1.4.3 Однокритериальные и многокритериальные задачи оптимизации

Выбор критерия эффективности функционирования для иерархических систем – всегда характеризуется многокритериальностью.

Для принятия решения в многокритериальной задаче оптимизации используется векторный показатель эффективности, который представляет собой некоторое математическое выражение в виде комбинации частных показателей. Такое выражение называют **критерием эффективности** или **целевой функцией**.

Необходимо иметь в виду, что векторный критерий вносит в задачу принятия решения неопределенность специального вида. Существуют различные подходы преодоления этой неопределенности, которые направлены на изыска-

ние путей отыскания дополнительной информации, позволяющей заменить векторный критерий его скалярным эквивалентом.

Одним из подходов является подход, формирующий так называемый обобщенный показатель эффективности (*мы его рассматривали как, аддитивный комплексный показатель в виде суммы частных показателей*).

Он представляет собой взвешенную сумму частных показателей W_i , в которую каждый из них входит с каким-то коэффициентом весомости α_i , отражающим его важность:

$$W = \alpha_1 W_1 + \alpha_2 W_2 + \dots + \alpha_n W_n ;$$

Такой подход имеет свои недостатки. Проблема заключается в обоснованном назначении коэффициентов весомости частных показателей эффективности.

Весовые коэффициенты можно определить, используя экспертное оценивание коэффициентов α_i . Экспертное оценивание может быть сделано, например, используя методы экспертных оценок – метод баллов или метод ранжирования (*ранговой корреляции*). В результате получим весовые коэффициенты для каждого показателя.

Существует еще один подход формирования решения - *метод последовательных уступок*. Он основан на сортировке показателей эффективности по степени важности в порядке убывания.

И, наконец, часто используемым подходом является *подход, выделяющий главный (доминирующий) показатель*. Среди множества показателей многокритериальной задачи выбирается один - «главный» показатель, а на остальные накладываются ограничения. При таком подходе все показатели, кроме главного, переводятся в разряд заданных условий, отвечающих реальным ресурсным возможностям ИС.

По сути, задача сводится к однокритериальной оптимизации, а критерий эффективности вводится на основе определенной концепции принятия решений человеком.

1.4.4 Множество Парето

Наиболее объективным и «научным», с точки зрения применения математического аппарата, является подход, формирующий так называемые «эффективные» решения. Суть такого подхода заключается в следующем.

Существует область компромисса, т.е. область не улучшаемых решений или так называемое множество Парето. Оно обладает тем свойством, что внутри него любой критерий не может быть улучшен без ухудшения хотя бы одного другого.

Предположим, что мы сделали некоторый выбор X . Далее предположим, что существует некоторый другой выбор X^* , такой, что для всех частных критериев $f_i(x)$ имеет место неравенство

$$f_i(x^*) \geq f_i(x) \quad i = 1, 2, \dots, n, \dots,$$

причем хотя бы одно из неравенств строгое.

Очевидно, в этом случае выбор X^* предпочтительнее выбора X . Поэтому все так называемые не эффективные оценки следует исключить из рассмотрения.

Таким образом, элемент $x^* \in X$ называется *оптимальным по Парето*, если не существует такого $x \in X$, который будет «лучше» x^* .

Смысл введенного понятия эффективного решения состоит в том, что оптимальное решение многокритериальной задачи следует искать *только* среди множества эффективных оценок. Это и есть принцип Парето.

Многомерный случай трудно представить наглядно, поэтому рассмотрим графический двумерный случай. Пусть имеется две критериальные функции (рис.1.4), определяющие цели системы $f_1(x)$ и $f_2(x)$, при этом для достижения наилучшего результата обе функции необходимо увеличить, т.е. $f_1(x) \rightarrow \max$; $f_2(x) \rightarrow \max$.

Возьмем плоскость в координатах (f_1 и f_2). Тогда каждому допустимому значению переменной x отвечает одна точка на плоскости (f_1 и f_2). Равенства $f_1 = f_1(x)$ и $f_2 = f_2(x)$ определяют параметрическое задание некоторой кривой $abcd$ на этой плоскости.

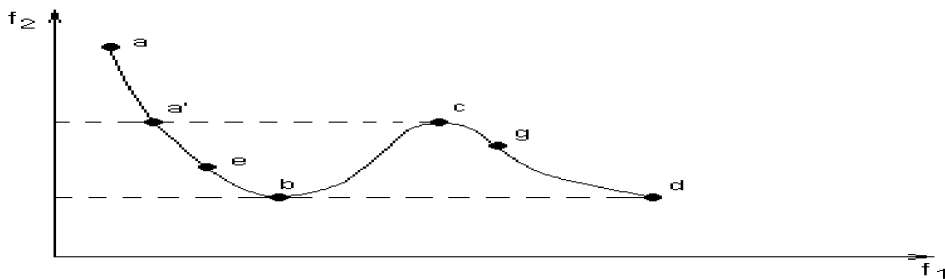


Рис.1.4 Графический двумерный случай

Рассмотрим, к какому участку обозначенной кривой можно применить принцип Парето.

Так, участок bc не может быть отнесён к множеству Парето, т.к. с ростом f_1 растёт и f_2 . На этом участке изменению переменной x соответствует одновременное увеличение обеих целевых функций, следовательно, эти варианты следует исключить из рассмотрения. Из тех же соображений исключаем и участок $a'b$.

На этом участке любой точке e на участке cd найдётся некоторая точка g , для которой и $f_{1g} > f_{1e}$ и $f_{2g} > f_{2e}$

На изображенном графике претендовать на принадлежность к множеству Парето могут участки cd и участок aa' (без точки a').

Таким образом, принцип Парето заключается в том, чтобы в качестве решения принимались те значения x (тот вектор x), которые принадлежат множеству Парето, т.е. удовлетворяют условию $f_i(x') \geq f_i(x)$. $i = 1, 2, \dots, n$.

Принцип Парето не выделяет одного единственного решения, а указывает на некоторое множество решений, сужая, таким образом, множество возможных альтернатив. В пределах этого выделенного множества делается окончательный выбор.

Лекция № 2.1

Раздел 2. Анализ систем методами исследования операций

Тема № 2.1 Анализ систем по схеме марковских случайных процессов

2.1.1. Марковские случайные процессы.

2.1.2. Марковская цепь.

2.1.3. Марковские процессы с дискретным состоянием и непрерывным временем.

2.1.4. Полумарковские процессы

2.1.1 Марковские случайные процессы

Потоком событий называется последовательность однородных событий, появляющихся одно за другим в случайные моменты времени. Примеры: поток вызовов на телефонной станции; поток сбоев ЭВМ; поток заявок на проведение регламентных работ и т.п.

Для математического описания многих явлений, развивающихся в форме случайного процесса, может быть с успехом применен математический аппарат, разработанный в теории вероятностей для так называемых марковских случайных процессов.

Благодаря сравнительной простоте и наглядности математического аппарата, высокой достоверности и точности получаемых решений особое внимание марковские процессы приобрели у специалистов, занимающихся исследованием операций и теорией принятия оптимальных решений.

Случайный процесс называется марковским процессом (или процессом без последствия, или процессом без памяти), если для каждого момента времени t_0 вероятность любого состояния системы в будущем (*при* $t > t_0$) зависит только от её состояния в настоящем (*при* $t = t_0$) и не зависит от того, когда и каким образом система пришла в это состояние (*иначе: при фиксированном настоящем будущее не зависит от предистории процесса – от прошлого*).

Другими словами, в марковском СП будущее развитие зависит только от его настоящего состояния и не зависит от «предистории» процесса.

Или другое определение.

Процесс называют марковским, если состояние системы S_i в некоторый момент времени определяет лишь вероятность $p_{ij}(t)$ того, что через промежуток времени t система будет находиться в состоянии S_j , причём эта вероятность не зависит от течения процесса в предшествующий период.

Вероятности $p_{ij}(t)$ называют *переходными вероятностями*.

Виды марковских случайных процессов.

Марковские случайные процессы делятся на классы. Первый классификационный признак – характер спектра состояний.

Случайный процесс (СП) называется процессом с дискретными состояниями, если возможные состояния системы S_1, S_2, S_3, \dots можно перечислить, а сам процесс состоит в том, что время от времени система S скачком (*мгновенно*) перескакивает из одного состояния в другое. При этом моменты времени переходов являются случайными.

Существуют процессы с непрерывными состояниями (*плавный переход из состояния в состояние*). Пример такого процесса: изменение напряжения в осветительной сети.

Второй классификационный признак – характер функционирования во времени. СП называется процессом с дискретным временем, если переходы системы из состояния в состояние возможны только в строго определенные, заранее фиксированные моменты времени: t_1, t_2, \dots . Если переход системы из состояния в состояние возможен в любой наперед неизвестный случайный момент, то говорят о СП с непрерывным временем.

Различают следующую классификацию, применяемую к случайным марковским процессам [1].

- 1) марковские цепи;
- 2) марковские последовательности;
- 3) марковские процессы с конечным и бесконечным числом состояний;
- 4) дискретно-непрерывные марковские процессы;
- 5) смешанные марковские процессы.

Характер реализаций четырех основных видов марковских процессов приведен на рис. 2.1.



Рис. 2.1 Характер реализаций основных видов марковских процессов
Мы рассмотрим два варианта описания марковских процессов - с дискретным и непрерывным временем.

2.1.2 Марковская цепь

Марковской цепью называется марковский случайный процесс с дискретным состоянием и дискретным временем.

Случайный процесс называется процессом с *дискретными состояниями*, если возможные состояния системы S_1, S_2, S_3, \dots можно перечислить (*пронумеровать*) одно за другим, а сам процесс состоит в том, что время от времени система S скачком (*мгновенно*) переходит из одного состояния в другое.

Модель марковского процесса с дискретным состоянием и дискретным временем удобно представлять в виде графа рис. 2.2.

При этом - состояния (вершины) связаны между собой связями (*переходами из i -го состояния в j -е состояние*).

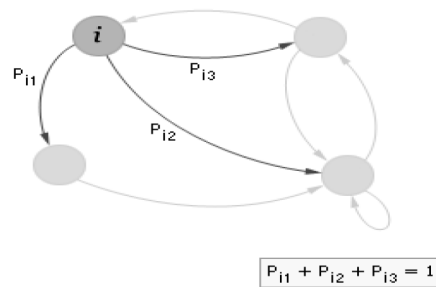


Рис. 2.2 Пример графа переходов

У стрелок проставляются соответствующие *вероятности переходов*. Такой граф состояний называют **размеченным**.

Для такого процесса моменты t_1, t_2, \dots , когда система S может менять свое состояние, удобно рассматривать как последовательные шаги процесса, а в качестве аргумента, от которого зависит процесс, рассматривать не время t , а номер шага: $1, 2, \dots, k, \dots$.

Случайный процесс в этом случае характеризуется последовательностью состояний

$$S(0), S(1), S(2), \dots, S(k), \dots, (*)$$

если $S(0)$ – начальное состояние системы (*перед первым шагом*); $S(1)$ - состояние системы непосредственно после первого шага; $\dots, S(k)$ – состояние системы непосредственно после k -го шага \dots .

Таким образом, переход осуществляется на каждом шаге, то есть исследователя интересует только последовательность состояний, которую проходит СП в своем развитии, и не интересует, когда конкретно происходил каждый из переходов.

Марковская цепь называется **однородной**, если переходные вероятности не зависят от номера шага, а зависят только от того, из какого состояния и в какое осуществляется переход, в противном случае - неоднородная МЦ.

Для однородной марковской цепи для n состояний системы переходные вероятности p_{ij} могут быть записаны в виде квадратной матрицы размером n на n .

$$\|P_{ij}\| = \begin{vmatrix} P_{11} & P_{12} \dots P_{1j} \dots P_{1n} \\ P_{21} & P_{22} \dots P_{2j} \dots P_{2n} \\ \dots & \dots & \dots & \dots \\ P_{i1} & P_{i2} \dots P_{ij} \dots P_{in} \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} \dots P_{nj} \dots P_{nn} \end{vmatrix}.$$

$$\sum_{j=1}^n P_{ij} = 1 \quad (i = 1, \dots, n).$$

Квадратная матрица, элементы которой неотрицательны и сумма элементов, стоящих в каждой строке (или столбце) равна единице, называется *стохастической* матрицей.

Если для марковской цепи заданы начальное распределение вероятностей $p_1(0), p_2(0), \dots, p_i(0), \dots, p_n(0)$ и матрица переходных вероятностей p_{ij} , то вероятности состояний системы могут быть определены по рекуррентной формуле

$$p_i(k) = \sum_{j=1}^n p_j(k-1) P_{ij} \quad (i = 1, \dots, n; j = 1, \dots, n).$$

При фактических вычислениях по указанной формуле надо в ней учитывать не все состояния S_i , а только те, для которых переходные вероятности отличны от нуля.

2.1.3 Марковские процессы с дискретным состоянием и непрерывным временем

На практике значительно чаще встречаются ситуации, когда переходы системы из состояния в состояние происходит в случайные моменты времени, которые заранее указать невозможно.

Для описания таких процессов может быть с успехом применена схема марковского случайного процесса с дискретными состояниями и непрерывным временем – *непрерывная цепь Маркова*.

Для такого процесса вероятность перехода из состояния S_i в S_j для любого фиксированного момента времени равна нулю, так же как вероятность любого отдельного значения НСВ.

Вместо вероятности перехода P_{ij} рассматривается плотность вероятности перехода λ_{ij} .

С точностью до бесконечно малых больших порядков можно показать, что:

$$P_{ij}(\Delta t) = \lambda_{ij} \cdot \Delta t;$$

При рассмотрении непрерывной цепи Маркова удобно представлять себе переходы системы из состояния в состояния как происходящие под влиянием некоторых потоков событий.

В этом случае плотности вероятностей перехода получают смысл интенсивностей λ_{ij} соответствующих потоков событий (как только происходит первое событие в потоке с интенсивностью λ_{ij} , система из состояния S_i скачком переходит в состояние S_j). Если все эти потоки пуассоновские, то процесс, протекающий в системе будет марковским.

Таким образом, для непрерывных цепей Маркова вместо переходных вероятностей p_{ij} у стрелок указываются плотность вероятностей переходов λ_{ij} (рис. 2.3).

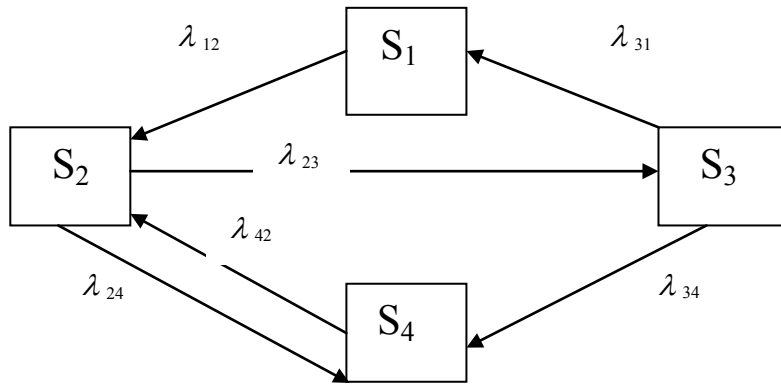


Рис. 2.3 Граф состояний системы

Если плотность вероятностей переходов λ_{ij} не зависит от времени t (от того, в какой момент начинается элементарный участок Δt), то такой марковский процесс называется *однородным*.

Рассмотрим случай, когда система S имеет конечное число состояний

$$S_1, S_2, \dots, S_n.$$

Для описания случайного процесса, протекающего в этой системе, применяются вероятности состояний

$$p_1(t), p_2(t), \dots, p_n(t), (*)$$

где $p_i(t)$ – вероятность того, что система S в момент t находится в состоянии S_i :

Очевидно, что любого t

$$\sum_{i=1}^n p_i(t) = 1.$$

Для нахождения вероятностей (*) необходимо решить систему дифференциальных уравнений (*уравнений Колмогорова*), имеющих вид:

$$\frac{dP_i(t)}{dt} = \sum_{j=1}^n \lambda_{ji}(t)P_j(t) - \sum_{j=1}^n \lambda_{ij}(t)P_i(t); \quad (i, j = 1, 2, \dots, n)$$

Величина $\lambda_{ij}P_i(t)$ называется *потоком вероятности перехода* из состояния S_i в состояние S_j .

Уравнения Колмогорова составляют по размеченному графу состояний системы, пользуясь следующим правилом:

- скорость изменения вероятности (*производная*) каждого состояния равна алгебраической сумме произведений интенсивности потока, переводящего систему по данному направлению, на вероятность того состояния, откуда осуществляется переход.

- мнемоническое правило:

- знак «+», если стрелка в данное состояние;
- знак «-», если стрелка из данного состояния.

Это правило составления дифференциальных уравнений для вероятностей состояний является общим и справедливо для любой непрерывной марковской цепи. С его помощью можно совершенно механически, без всяких рассуждений, записывать дифференциальные уравнения для вероятностей состояний непосредственно по размеченному графу состояний [7].

Для графа состояний рис. 2.5, уравнения имеют вид:

$$\begin{cases} \frac{dp_1(t)}{dt} = -\lambda_{12}p_1(t) + \lambda_{31}p_3(t); \\ \frac{dp_2(t)}{dt} = -\lambda_{23}p_2(t) - \lambda_{24}p_2(t) + \lambda_{12}p_1(t) + \lambda_{42}p_4(t); \\ \frac{dp_3(t)}{dt} = -\lambda_{31}p_3(t) - \lambda_{34}p_3(t) + \lambda_{23}p_2(t); \\ \frac{dp_4(t)}{dt} = -\lambda_{42}p_4(t) + \lambda_{24}p_2(t) + \lambda_{34}p_3(t). \end{cases}$$

Для решения системы уравнений Колмогорова необходимо задать начальное распределение вероятностей $P_0(0), P_1(0), \dots, P_n(0)$.

Как правило, за исключением особенно простых систем, решение возможно получить лишь численными методами.

Интегрирование этих уравнений при известном начальном состоянии системы даст искомые вероятности состояний как функции времени.

Возникает вопрос, как будет вести себя система при $t \rightarrow \infty$?

Существуют ли пределы функций $P_1(t), P_2(t), \dots, P_i(t), \dots, P_n(t)$?

Если эти пределы существуют, то соответствующие вероятности состояний называются предельными вероятностями состояний (или «финальными», т.е. конечными). Если предельные вероятности существуют, то в этом состоянии имеет место установившийся режим, для которого производные будут равны нулю.

Система, для которой существуют финальные вероятности, называется эргодической и соответствующий случайный процесс – эргодическим.

В этом случае система дифференциальных уравнений Колмогорова превращается в систему алгебраических уравнений.

Совместно с нормирующим условием $\sum_{i=1}^n P_i(t) = 1$ эти уравнения дают воз-

можность вычислить все предельные вероятности состояний $P_1, P_2, \dots, P_i, \dots, P_n$.

В этом предельном режиме каждая финальная вероятность может быть истолкована как среднее относительное время пребывания системы в данном состоянии.

2.1.4 Полумарковские процессы

Теория полумарковских процессов (ПМП) является одним из интенсивно развиваемых направлений в теории случайных процессов.

Рассмотрим процесс с дискретными состояниями и непрерывным временем, схема которого изображена на рис. 2.4.

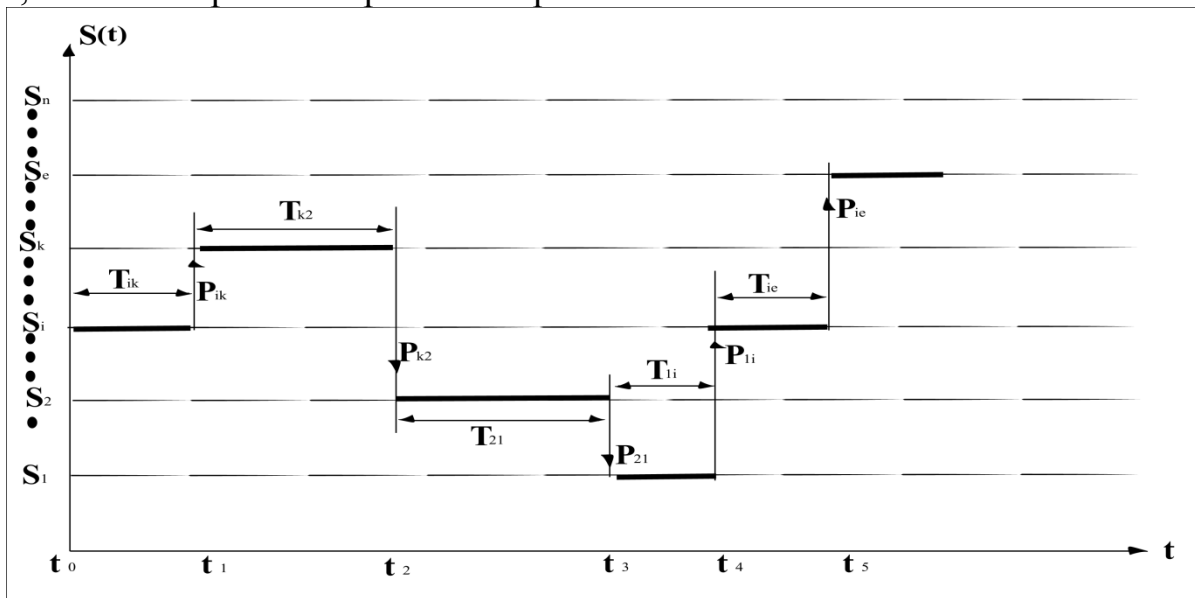


Рис. 2.4. Процесс с дискретными состояниями и непрерывным временем

$S_1, S_2, \dots, S_i, \dots, S_k, \dots, S_l, \dots, S_n$ - состояния системы. Исходное состояние - S_i

На схеме изображена следующая последовательность состояний:

$$S_i \rightarrow S_k \rightarrow S_2 \rightarrow S_1 \rightarrow S_i \rightarrow S_L \rightarrow \dots$$

Следует иметь в виду, что в данном случае это одна из возможных последовательностей реализации процесса.

Одношаговые вероятности переходов - суть $P_{ik}, P_{k2}, P_{21}, P_{1i}, P_{iL}, \dots$

Соответственно – T_{ij} - время ожидания в состоянии i до перехода в состояние j . Каждое из этих времен ожидания имеет какой-то закон распределения.

Если игнорировать случайный характер времен ожидания и рассматривать только моменты перехода, то процесс $S(t)$ будут представлять собой однородную цепь Маркова. Если же учитывать пребывание процесса в разных состояниях в течение случайного отрезка времени, то здесь возможны, в зависимости от характера распределения этого времени, два различных характера процесса.

Если распределение времени ожидания является экспоненциальным, т.е. поток, переводящий систему из состояния в состояние, является пуассоновским, то имеет место чисто марковский процесс с дискретными состояниями и непрерывным временем. В этом случае справедливы все уравнения - и дифференциальные и алгебраические для предельных вероятностей состояний.

Если же распределение времени ожидания является любым другим (*кроме экспоненциального*), в том числе это время может быть и некоторым постоянным числом, то процесс не является марковским в чистом виде.

Марковским процесс является только и только в момент перехода, при этом вероятности переходов между состояниями определяются матрицей вероятностей, как и для случая марковской цепи.

Случайный процесс, при котором переходы между состояниями являются марковскими, а времена нахождения в любом из состояний описываются произвольной функцией распределения, называется *полумарковским*. В литературе применяются также термины *вложенная цепь Маркова* или *вложенный марковский процесс*.

Смысл этих терминов состоит в том, что Марковский процесс перехода между состояниями системы происходит внутри другого процесса (*не Марковского*).

Для описания полумарковского процесса необходимо задать:

- граф состояний и переходов, т.е. все возможные состояния $S_i (i = 1, 2, \dots, n)$ и все возможные переходы внутри графа с указанием их направлений, что эквивалентно заданию матрицы переходных вероятностей P_{ij} ;

- матрицу $Q = \{Q_{ij}(t)\}$ независимых функцией распределения времени пребывания в i -м состоянии перед переходом в j -е состояние;

- начальное состояние в момент $t = 0$.

На основании этих исходных данных определяются вероятности P_{ij} перехода из состояния i в состояние j в момент скачка и математические ожидания \bar{t}_i времени пребывания в i -м состоянии.

Лекция №2.2

Раздел 2. Анализ систем методами исследования операций

Тема № 2.2 Анализ систем массового обслуживания

- 2.2.1. Компоненты систем массового обслуживания.
- 2.2.2. Классификация систем массового обслуживания.
- 2.2.3. Показатели качества обслуживания СМО.
- 2.2.4. Анализ систем массового обслуживания с отказами.
 - Система $M/M/1/0$ – одноканальная СМО с отказами.
 - Система $M/M/n/0$ – многоканальная СМО с отказами.
- 2.2.5. Анализ систем массового обслуживания с ожиданием.
 - Системы $M/M/1/m$ и $M/M/1/\infty$ – одноканальные системы с ожиданием.
 - Система $M/M/n/m$ и $M/M/n/\infty$ – многоканальные системы с ожиданием.

2.2.1 Компоненты систем массового обслуживания

Большинство систем, с которыми человек имеет дело, являются стохастическими. Попытка их математического описания с помощью детерминистических моделей приводит к огрублению истинного положения вещей. При решении задач анализа и проектирования таких систем приходится считаться с положением вещей, когда случайность является определяющей для процессов, протекающих в системах. При этом пренебрежение случайностью, попытка «втиснуть» решение перечисленных задач в детерминистические рамки приводит к искажению, к ошибкам в выводах и практических рекомендациях.

Первые задачи теории систем массового обслуживания (ТСМО) были рассмотрены сотрудником Копенгагенской телефонной компании, датским ученым А.К. Эрлангом (1878- 1929гг) в период между 1908 и 1922гг.

Эти задачи были вызваны к жизни стремлением упорядочить работу телефонной сети и разработать методы, позволяющие заранее повысить качество обслуживания потребителей в зависимости от числа используемых устройств. Оказалось, что ситуации, возникающие на телефонных станциях, являются типичными не только для телефонной связи. Работа аэродромов, морских и речных портов, магазинов и т.д. может быть описана в рамках ТСМО.

Следует сказать и о терминах «ТМО» и «теория очередей».

Во многих работах они трактуются как равнозначные, в других — теория очередей рассматривается лишь как раздел ТМО, поскольку последней изучаются системы не только с очередями, но и с отказами (напр., когда телефонная станция занята, очередь абонентов не образуется), а также некоторые иные.

Типовую систему массового обслуживания (СМО) можно изобразить в виде (рис.2.5).



Рис. 2.5 Типовая СМО

Основными компонентами СМО любого вида являются.

1. Входящий поток заявок, который может быть охарактеризован интенсивностью потока λ – среднее число заявок, поступающих в единицу времени.

2. Приборы (*каналы*) обслуживания, которых в системе может быть один (*одноканальная система*) или несколько (*многоканальная система*).

Роль каналов могут играть различные приборы, лица, выполняющие те или иные операции (*кассиры, операторы, парикмахеры, продавцы*), линии связи, автомашины, краны, ремонтные бригады, бензоколонки и т.д.

3. Накопители (*устройства для обеспечения ожидания обслуживания*), которые могут располагаться как перед всей системой, так и перед каждым каналом обслуживания.

4. Выходящий поток обслуженных заявок, который может быть охарактеризован интенсивностью обслуживания μ - среднее число заявок, выполняемых в единицу времени.

Структура СМО определяется количеством и типом обслуживающих приборов, а также наличием накопителей.

Характер обслуживания состоит в порядке обслуживания, который принято называть дисциплиной обслуживания, и величиной μ - интенсивностью обслуживания. Выделяются следующие виды дисциплины обслуживания:

- первым пришел - первый обслуживаешься;
- пришел последним - обслуживаешься первым;
- случайный отбор заявок;
- отбор заявок по критерию приоритетности;
- ограничение времени ожидания момента наступления обслуживания.

Характерной особенностью задач СМО являются условия «двойной» случайности:

- случаен момент времени поступления заявки на обслуживание;
- случайная длительность времени этого обслуживания.

Случайный характер потока заявок и длительности их обслуживания порождает в СМО случайный процесс. Случайным процессом (*или случайной функцией*) называется соответствие, при котором каждому значению аргумента (*в данном случае - моменту из промежутка времени проводимого опыта*) ставится в соответствие случайная величина (*в данном случае - состояние СМО*).

2.2.2 Классификация систем массового обслуживания

Существует несколько подходов к классификации СМО.

Так, например. По характеру источника требований или как иногда говорят, по кругу обслуживаемых объектов, различают СМО с конечным и бесконечным количеством заявок (*требований*) на входе.

В первом случае в системе циркулирует конечное, обычно постоянное количество заявок, которые после завершения обслуживания возвращаются в источник. Во втором случае источник генерирует бесконечное число заявок. Первый вид СМО называют замкнутой, второй – разомкнутой.

Наиболее распространенной и общепринятой классификацией является классификация, введенная английским математиком *Кендаллом*.

По его классификации, любая СМО может быть представлена четырьмя параметрами в виде следующей записи

$$P_{ВХ} / B_{об} / N_{пр} / E_{нак}$$

$P_{ВХ}$ - характер входящего потока требований,

$B_{об}$ - распределение времени обслуживания;

$N_{пр}$ - число обслуживающих приборов;

$E_{нак}$ - емкость накопителя (длина очереди).

Характер входящего потока требований принято обозначать следующими символами:

- M (*Markovian*) – входящий поток требований является Пуассоновским, т.е. распределение времени между поступающими заявками подчинено экспоненциальному закону;

- E (*Erlangian*) – входящий поток является Эрланговским;

- D (*Deterministic*) – детерминированный постоянный поток;

- G (*General*) – произвольный рекуррентный поток.

Те же символы применяются и для обозначения распределения времени обслуживания:

- M - распределение по экспоненциальному закону;

- E - распределение по закону Эрланга;

- D - время обслуживания постоянная величина;

- G - произвольное распределение времени обслуживания.

Число обслуживающих приборов – равно или больше единицы.

При $N_{пр} = 1$ – одноканальная СМО;

При $N_{пр} > 1$ – многоканальная СМО.

Емкость накопителя может варьироваться от 0 до ∞ .

При $E_{нак} = 0$ поступившая заявка в случае, если все каналы заняты, теряется (*получает отказ в обслуживании*). Такие СМО принято называть – системы с потерями (*отказами*).

При $E_{\text{нак}} > 1$ система является системой с ожиданием (*с очередью*). В таких системах заявка, поступившая в момент, когда все каналы заняты, отправляется в накопитель (*становится в очередь*) и ожидает освобождения хотя бы одного канала. Системы массового обслуживания с ожиданием делятся на системы с ограниченным ожиданием и системы с неограниченным ожиданием.

В системах с ограниченным ожиданием может ограничиваться:

- длина очереди;
- время пребывания в очереди.

В системах с неограниченным ожиданием заявка, стоящая в очереди, ждет обслуживание неограниченно долго, т.е. пока не подойдет очередь.

В системах, в которых входящий поток является пуассоновским, а распределение времени обслуживания – экспоненциальное, процессы обслуживания являются Марковскими процессами.

Анализ таких систем может быть сделан методами, изложенными ранее.

Приведем условное обозначение наиболее распространенных систем:

- М/М/1/0 - одноканальная СМО с отказами;
- М/М/n/0 - многоканальная СМО с отказами;
- М/М/1/m - одноканальная СМО с ожиданием ($E_{\text{нак}} = m$);
- М/М/n/m - многоканальная СМО с ожиданием, но с возможностью отказа (*число каналов – n, $E_{\text{нак}} = m$*);
- М/М/1/∞ - одноканальная СМО с ожиданием без отказа ($E_{\text{нак}} = \infty$).

Приведенная классификация СМО является условной. На практике чаще всего СМО выступают в качестве смешанных систем. Например, заявки ожидают начала обслуживания до определенного момента, после чего система начинает работать как система с отказами.

2.2.3 Показатели качества обслуживания СМО

Каждая СМО в зависимости от своих параметров - характера потока заявок, числа каналов обслуживания и их производительности, а также от правил организации работы обладает определенной эффективностью функционирования (*пропускной способностью*), позволяющей ей более или менее успешно справляться с потоком заявок.

В зависимости от типа СМО при оценке качества ее функционирования могут применяться различные показатели.

Рассмотрим основные показатели.

1. $P_{\text{отк}}$ - вероятность отказа заявке (*потери заявки*) в обслуживании или доля из общего числа требований, которым будет отказано в обслуживании.

Заявка может получить отказ из-за занятости всех каналов и мест ожидания (*если есть накопитель*) или из-за превышения максимального времени ожидания.

2. P_0 – вероятность того, что поступившая заявка немедленно будет принята к обслуживанию, другими словами – это есть то, что СМО свободна.

Еще одно название – доля времени, когда все каналы свободны (*вероятность простоя*). Является промежуточной величиной расчета всех СМО. Через нее выражаются многие остальные параметры.

3. A - абсолютная пропускная способность (*абсолютная эффективность обслуживания*). Это среднее число заявок, которое может обслужить СМО в единицу времени, т.е. число обслуженных за день, за час, за минуту заявок, клиентов и т.п.

4. q - относительная пропускная способность - отношение среднего числа заявок, обслуживаемых СМО в единицу времени, к среднему числу поступивших заявок за это же время.

Еще одно название – средняя доля поступивших заявок, обслуживаемая системой (*или вероятность обслуживания поступившей заявки*).

$$q = 1 - P_{\text{отк}}$$

Величины A и q связаны соотношением

$$A = \lambda q.$$

Для СМО с неограниченным ожиданием каждая поступившая заявка рано или поздно будет обслужена, поэтому $q = 1$ и $A = \lambda$.

5. ρ - коэффициент загрузки СМО или приведенная интенсивность потока заявок

$$\rho = \frac{\lambda}{\mu}$$

Другими словами, величина ρ представляет собой среднее число каналов, которое должно быть для обслуживания в единицу времени всех поступающих заявок. До ее вычисления параметры λ и μ обязательно должны быть приведены к единой единице времени!

6. \bar{Z} - среднее число занятых каналов обслуживания (*число работающих мастеров, занятых телефонов и т.п. для многоканальной системы*).

7. \bar{W} - среднее число заявок под обслуживанием.

Для многоканальной системы без очереди или с ограниченной длиной очереди – совпадает со средним значением занятых каналов.

8. $\bar{t}_{\text{ож}}$ - среднее время ожидания заявки в очереди.

9. $\bar{t}_{\text{сист}}$ - среднее время пребывания заявки в СМО (*в очереди и под обслуживанием*).

10. \bar{r} - среднее число заявок, находящихся в очереди, т.е. длина очереди.

11. \bar{k} - среднее число заявок, находящихся в СМО.

Важно понимать, что все эти величины носят средне статистический характер, то есть реализуются «в среднем» при продолжительном времени устойчивой работы системы массового обслуживания.

2.2.4 Анализ систем массового обслуживания с отказами

Независимо от характера процесса, протекающего в системе массового обслуживания, различают два основных вида систем массового обслуживания:

- системы с отказами, в которых заявка, поступившая в систему в момент, когда все каналы заняты, получает отказ и сразу же покидает очередь;
- системы с ожиданием (*очередью*), в которых заявка, поступившая в момент, когда все каналы обслуживания заняты, становится в очередь и ждет, пока не освободится один из каналов. Системы массового обслуживания с ожиданием делятся на системы с ограниченным ожиданием и системы с неограниченным ожиданием.

2.2.4.1 Система M/M/1/0 – одноканальная СМО с отказами

Рассмотрим простейшую из всех задач ТМО - задачу о функционировании одноканальной СМО с отказами.

В системе рассматриваемого типа поток заявок является пуассоновским с интенсивностью λ (среднее число заявок, поступающих в единицу времени).

$$\lambda = 1/t_{cp}$$

Обслуживание заявки продолжается в течение случайного времени $t_{обсл}$, распределенного также по показательному закону с параметром μ (среднее число заявок, обслуживаемых в единицу времени или иначе - интенсивность освобождения канала).

$$\mu = 1/t_{обсл}$$

Требуется найти:

- абсолютную (A) и относительную (q) пропускные способности.

Возможные состояния системы - всего два. Граф состояний приведен на рис. 2.6

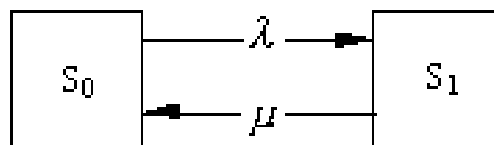


Рис. 2.6 Одноканальная СМО с отказами

S_0 - канал свободен и S_1 - канал занят.

Из состояния S_0 в S_1 систему переводит поток заявок с интенсивностью λ ; из S_1 в S_0 - «поток обслуживаний» с интенсивностью μ .

Обозначим вероятности состояний $P_0(t)$ и $P_1(t)$. Очевидно, для любого момента t - $P_0(t) + P_1(t) = 1$

Составим дифференциальные уравнения Колмогорова для вероятностей состояний.

Имеем:

$$\begin{cases} \frac{dP_0}{dt} = -\lambda P_0 + \mu P_1; \\ \frac{dP_1}{dt} = -\mu P_1 + \lambda P_0; \end{cases}$$

Решение этой системы уравнений дает следующие результаты:

$$q = \frac{\mu}{\lambda + \mu} \quad A = \lambda q = \lambda \frac{\mu}{\lambda + \mu}.$$

Зная q (вероятность того, что пришедшая в момент t заявка будет обслужена) легко найти вероятность отказа:

$$P_{отк} = 1 - q$$

$P_{отк}$ есть не что иное, как средняя доля необслуженных заявок среди поданных. В пределе

$$P_{отк} = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}$$

2.2.4.2 Система M/M/n/0 – многоканальная СМО с отказами

Состояния n -канальной системы.

- S_0 - система полностью свободна;
- S_1 - один канал занят, остальные каналы свободны;
- S_2 - занято два канала, остальные каналы свободны;
-
- S_i - занято i каналов, остальные каналы свободны;
-
- S_n - все n - каналов заняты.

Размеченный граф состояний системы представлен на рис. 2.7.

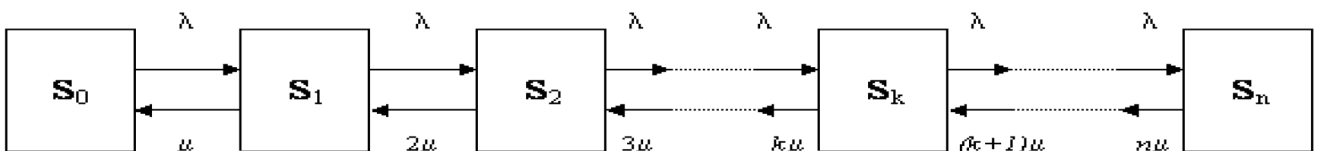


Рис.2.7 Граф состояний - многоканальная СМО с отказами

Вероятности состояний обозначаются соответственно $- P_0, P_1, P_2, \dots, P_n$.

Известно, что интенсивность перехода «слева направо» определяется λ .

Интенсивность перехода «справа налево» имеет свои особенности. Действительно, если СМО находится в состоянии S_n , то возврат в состояние S_{n-1} может произойти многими способами (*освобождается или первый, или второй, или ... канал*), что увеличивает интенсивность освобождения каналов до $n\mu$.

Таким образом, для переходов $S_{n-1}, S_{n-2}, \dots, S_2, S_1$, для которых μ заменяется на $(n-1)\mu, (n-2)\mu, \dots, 2\mu$ соответственно. Именно в этом проявляется эффект многоканальности, приводящий к убыстрению обслуживания в среднестатистическом смысле.

Применительно к данному графу состояний опишем процесс функционирования ИС системой дифференциальных уравнений:

$$\begin{cases} \frac{dp_0}{dt} = -\lambda p_0 + \mu p_1; \\ \frac{dp_1}{dt} = -(\lambda + \mu)p_1 + \lambda p_0 + 2\mu p_2; \\ \frac{dp_2}{dt} = -(\lambda + 2\mu)p_2 + \lambda p_1 + 3\mu p_3; \\ \dots \\ \frac{dp_k}{dt} = -(\lambda + k\mu)p_k + \lambda p_{k-1} + (k+1)\mu p_{k+1}; \\ \dots \\ \frac{dp_n}{dt} = -n\mu p_n + \lambda p_{n-1}. \end{cases}$$

Учтем нормирующее условие $-\sum_{k=0}^n P_k = 1$.

Интегрирование уравнений дает *формулы Эрланга* для вычисления искоемых величин. Практическое значение имеют предельные значения вероятности состояний

$$P_0, P_1, P_2, \dots, P_k,$$

Для этого систему дифференциальных уравнений преобразуем в систему алгебраических и получим.

$$P_0 = \frac{1}{1 + \frac{\lambda/\mu}{1!} + \frac{(\lambda/\mu)^2}{2!} + \dots + \frac{(\lambda/\mu)^n}{n!}};$$

$$P_k = \frac{\lambda^k}{\mu \circ 2\mu \circ \dots \circ k\mu} P_0 = \frac{(\lambda/\mu)^k}{k!} P_0;$$

Обозначим $\rho = \frac{\lambda}{\mu}$.

Величина ρ определена нами ранее.

С учетом этого обозначения, вышеприведенные соотношения принимают

вид:

$$P_0 = \frac{1}{1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!}} = \left[1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!} \right]^{-1};$$

$$P_k = \frac{\rho^k}{k!} P_0;$$

Имея вероятности состояний $P_0, P_1, P_2, \dots, P_k$, можно найти интересующие нас характеристики эффективности СМО.

1. Вероятность отказа.

Заявка получает отказ, если приходит в момент, когда все n каналов заняты.

$$P_{отк} = P_n = \frac{\rho^n}{n!} P_0$$

2. Относительная пропускная способность.

Вероятность того, что заявка будет принята к обслуживанию, дополняет $P_{отк}$ до единицы:

$$q = 1 - P_{отк}$$

3. Абсолютная пропускная способность:

$$A = \lambda q = \lambda(1 - P_{отк})$$

4. Среднее число заявок в системе - $\bar{k} = \bar{z}$ (здесь они совпадают).

Среднее число занятых каналов \bar{z} можно определить через абсолютную пропускную способность A , которая уже известна. Действительно, A есть не что иное, как среднее число заявок, обслуживаемых в единицу времени.

Один занятый канал обслуживает в среднем за единицу времени μ заявок; среднее число занятых каналов получится делением A на μ :

$$\bar{z} = \frac{A}{\mu} = \frac{\lambda(1 - P_{отк})}{\mu} = \rho(1 - P_{отк}).$$

2.2.5 Анализ систем массового обслуживания с ожиданием

2.2.5.1 Системы M/M/1/m и M/M/1/∞ – одноканальные системы с ожиданием

Рассмотрим сначала систему M/M/1/m – одноканальную с ограниченной очередью.

Возможные состояния этой системы.

- S_0 - канал свободен;
- S_1 - канал занят, очереди нет;
- S_2 - канал занят, одна заявка стоит в очереди;
-
- S_i - канал занят, $i-1$ заявок в очереди;
-
- S_{m+1} - канал занят, m заявок стоит в очереди (*накопитель полностью загружен*).

Заявка, приходящая в момент, когда система находится в состоянии S_{m+1} , получает отказ.

Граф состояний имеет вид (рис. 2.8):

Очередует

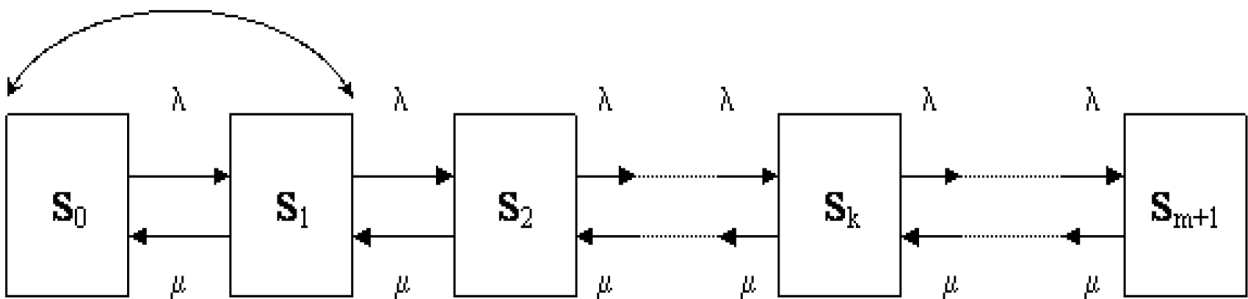


Рис.2.8 Граф состояний - одноканальная система с ожиданием

Для предельных вероятностей состояний можно записать:

$$\left\{ \begin{array}{l} p_1 = (\lambda / \mu) p_0 \\ p_2 = (\lambda / \mu)^2 p_0 \\ \dots \\ p_k = (\lambda / \mu)^k p_0 \\ \dots \\ p_{m+1} = (\lambda / \mu)^{m+1} p_0 \\ p_0 = \frac{1}{1 + (\lambda / \mu) + (\lambda / \mu)^2 + \dots + (\lambda / \mu)^{m+1}} \end{array} \right. \quad (*)$$

В знаменателе выражения для P_0 стоит геометрическая прогрессия, суммируя её находим

$$P_0 = \frac{1}{(1 - \rho^{m+2}) / (1 - \rho)} = \frac{1 - \rho}{1 - \rho^{m+2}}.$$

Это выражение справедливо только при $\rho < 1$ (при $\rho = 1$ она дает неопределенность вида $0/0$).

Имея вероятности состояний $P_0, P_1, P_2, \dots, P_k$, можно найти интересующие нас характеристики эффективности СМО.

1. Вероятность отказа.

Очевидно, заявка получает отказ только в случае, когда канал занят и все m мест в очереди тоже:

$$P_{отк} = P_{m+1} = \frac{\rho^{m+1}(1-\rho)}{1-\rho^{m+2}}$$

2. Относительная пропускная способность:

$$q = 1 - P_{отк} = 1 - \frac{\rho^{m+1}(1-\rho)}{1-\rho^{m+2}}$$

3. Абсолютная пропускная способность:

$$A = \lambda q$$

4. Среднее число занятых каналов - \bar{z} - не задействован.

5. Средняя длина очереди.

$$\bar{r} = \frac{\rho^2 \left[1 - \rho^m (m+1 - m\rho) \right]}{(1-\rho^{m+2})(1-\rho)}$$

6. Среднее число заявок, находящихся под обслуживанием.

Поскольку канал один, число обслуживаемых заявок может равняться 0 (с вероятностью P_0) или 1 (с вероятностью $1 - P_0$), откуда:

$$\varpi = 0 \circ P_0 + 1 \circ (1 - P_0) = \frac{\rho - \rho^{m+2}}{1 - \rho^{m+2}}$$

7. Среднее число заявок, находящихся в системе (как стоящих в очереди, так и находящихся на обслуживании).

Поскольку $\bar{k} = \bar{r} + \varpi$,

$$\bar{k} = \bar{r} + \frac{\rho - \rho^{m+2}}{1 - \rho^{m+2}}$$

8. Среднее время ожидания заявки в очереди.

Оно очевидно равно средней длине очереди, умноженной на среднее значение времени в потоке заявок. Для пуассоновского потока заявок среднее время $t_{cp} = 1/\lambda$.

Таким образом, имеем

$$\bar{t}_{OЖ} = \frac{\bar{r}}{\lambda}$$

9. Среднее время пребывания заявки в системе.

Оно складывается из среднего времени ожидания в очереди $t_{OЖ}$ и среднего времени обслуживания $t_{обсл}$.

Если загрузка системы составляет 100 %, очевидно, $t_{обсл} = 1/\mu$, в противном же случае

$$t_{обсл} = q/\mu.$$

Отсюда

$$\bar{t}_{СИСТ} = \frac{\bar{r}}{\lambda} + \frac{q}{\mu}$$

Для системы $M/M/1/\infty$, т.е. одноканальной системы с бесконечной очередью, установившийся режим возможен только при $\rho < 1$. При $\rho > 1$ очередь растёт до бесконечности.

Вероятности состояний системы $M/M/1/\infty$ для случая $\rho < 1$ можно получить из выражений для системы $M/M/1/m$ путем предельного перехода $m \rightarrow \infty$.

Эти выражения принимают вид:

$$\left. \begin{array}{l} P_0 = 1 - \rho \\ P_1 = \rho \cdot (1 - \rho) \\ P_2 = \rho^2 \cdot (1 - \rho) \\ \dots\dots\dots \\ P_k = \rho^k \cdot (1 - \rho) \\ \dots\dots\dots \end{array} \right\}$$

При отсутствии ограничений по длине очереди каждая заявка, пришедшая в систему, будет обслужена, поэтому $q = 1$,

$$A = \lambda q = \lambda$$

1. Среднее число заявок в очереди:

$$\bar{r} = \frac{\rho^2}{(1 - \rho)}$$

2. Среднее число заявок в системе:

$$\bar{k} = \bar{r} + \rho = \frac{\rho^2}{1 - \rho} + \rho = \frac{\rho}{1 - \rho}$$

3. Среднее время ожидания:

$$\bar{t}_{OЖ} = \frac{1}{\mu} \frac{\rho}{(1 - \rho)} = \frac{\rho^2}{\lambda(1 - \rho)}$$

4. Среднее время пребывания заявки в СМО:

$$\bar{t}_{СИСТ} = \bar{t}_{OЖ} + t_{обсл} = \frac{1}{\mu} \frac{1}{(1 - \rho)}$$

Расчетные формулы основных характеристик СМО типов $M/M/1/m$ и $M/M/1/\infty$ ($\rho < 1$) представлены в табл. 5.4 Пособия по проведению практических занятий (изд. МГТУ ГА, 2013).

2.2.5.2 Система $M/M/n/m$ – многоканальная система с ожиданием

Рассмотрим n - канальную СМО с ожиданием, на которую поступает поток заявок с интенсивностью λ .

Интенсивность обслуживания - одного канала μ ; число мест в очереди m .

Состояния системы нумеруются по числу заявок, связанных системой:

нет очереди:

S_0 - все каналы свободны;

S_1 - занят один канал, остальные свободны;

S_k - заняты k каналов, остальные нет;

S_n - заняты все n каналов, свободных нет;

есть очередь:

S_{n+1} - заняты все n каналов; одна заявка стоит в очереди;

S_{n+r} - заняты все n каналов, r заявок в очереди;

S_{n+m} - заняты все n каналов, m заявок в очереди.

Граф состояний имеет вид (рис. 2.9):

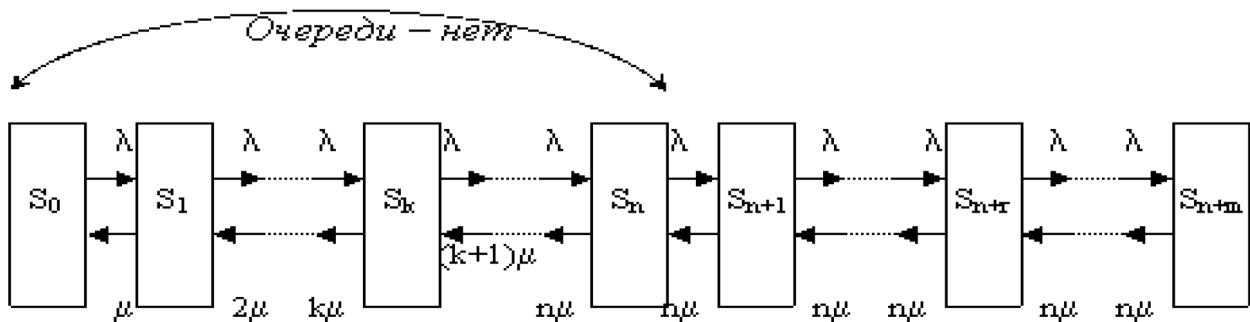


Рис. 2.9 Многоканальная СМО с ожиданием

Выражения для предельных вероятностей

$$\begin{cases} p_k = \frac{\rho^k}{k!} p_0, & (k = 1, 2, \dots, n); \\ p_{n+i} = \frac{\rho^{n+i}}{n^i n!} p_0, & (i = 1, 2, \dots, m); \\ p_0 = \left(1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n n!} + \frac{\rho^{n+2}}{n^2 n!} + \dots + \frac{\rho^{n+m}}{n^m n!} \right)^{-1} = \\ = \left(1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!} + \frac{\rho^n}{n!} \cdot \frac{\rho - \left(\frac{\rho}{n}\right)^{m+1}}{1 - \frac{\rho}{n}} \right)^{-1}. \end{cases}$$

Получив выражения для предельных вероятностей можно вычислить вероятности состояний $P_0, P_1, P_2, \dots, P_k$, и найти интересующие нас характеристики эффективности СМО.

Расчетные формулы основных характеристик СМО типов $M/M/n/m$ и $M/M/n/\infty$ представлены в табл. 5.4 Пособия по проведению практических занятий (изд. МГТУ ГА, 2013).

1. Вероятность отказа.

Поступившая заявка получает отказ, если заняты все n каналов и все m

мест в очереди:

$$P_{\text{ОТК}} = P_{n+m} = \frac{\rho^{n+m}}{n^m n!} P_0$$

2. **Относительная пропускная способность** дополняет вероятность отказа до единицы:

$$q = 1 - P_{\text{ОТК}} = 1 - \frac{\rho^{n+m}}{n^m n!} P_0$$

3. **Абсолютная пропускная способность** СМО:

$$A = \lambda q = \lambda \left(1 - \frac{\rho^{n+m}}{n^m n!} P_0 \right)$$

4. **Среднее число занятых каналов.**

$$\bar{z} = \frac{A}{\mu} = \frac{\lambda}{\mu} \left(1 - \frac{\rho^{n+m}}{n^m n!} P_0 \right) = \rho \left(1 - \frac{\rho^{n+m}}{n^m n!} P_0 \right).$$

5. **Среднее число заявок в очереди** можно вычислить непосредственно как математическое ожидание дискретной случайной величины:

$$\bar{r} = \frac{\rho^{n+1}}{n n!} P_0 \left(1 + 2\chi + 3\chi^2 + \dots + m\chi^{m-1} \right), \quad \text{где } \chi = \frac{\rho}{n}.$$

Здесь (выражение в скобках) встречается производная суммы геометрической прогрессии, используя соотношение для нее, получаем:

$$\bar{r} = \frac{\rho^{n+1}}{n! n} \cdot \frac{1 - \left(\frac{\rho}{n}\right)^m \left(m + 1 + m \frac{\rho}{n}\right)}{\left(1 - \frac{\rho}{n}\right)^2} P_0$$

6. Среднее число заявок в системе:

$$\bar{k} = \bar{r} + \bar{z}$$

7. Среднее время ожидания заявки в очереди.

$$\bar{t}_{ож} = \frac{\bar{r}}{\lambda} = \frac{\rho^n P_0}{n\mu!} \frac{1 - \chi^m (m + 1 - m\chi)}{(1 - \chi)^2}$$

8. Среднее время пребывания заявки в системе, так же, как и для одно-канальной СМО, отличается от среднего времени ожидания на среднее время обслуживания, умноженное на относительную пропускную способность:

$$\bar{t}_{СМО} = \bar{t}_{ож} + \bar{t}_{обсл} = \bar{t}_{ож} + \frac{q}{\mu}.$$

Выражения для параметров системы типа М/М/п/∞ можно получить из формул путём предельного перехода $m \rightarrow \infty$.

Полученные ниже выражения справедливы только при $\frac{\rho}{n} < 1$.

Лекция № 2.3

Раздел 2. Анализ систем методами исследования операций

Тема № 2.3 Метод сетевого планирования

2.3.1 Классификация методов планирования работы.

2.3.2 Сетевой метод планирования работы.

2.3.3 Оптимизация сетевых графиков.

2.3.1 Классификация методов планирования работы

Планирование – одна из функций управления. Его ключевая роль определяется тем, что планирование предшествует выполнению всех других управленческих функций, поскольку оно призвано формулировать цели и задачи организации, пути и методы их достижения, а также определять необходимые для этого средства.

Посредством планирования руководство стремится установить основные направления усилий и принятия решений, которые обеспечат единство цели для всех членов организации. Другими словами, планирование - это один из способов, с помощью которого руководство обеспечивает единое направление усилий всех членов организации к достижению ее общих целей.

Планирование работы подразделяется на:

- **общее планирование**, т.е. планирование на календарные сроки. К числу календарных планов относятся: годовые, месячные и ежедневные планы работы;
- **технологическое планирование** предусматривает разработку типовых технологических графиков видов работ по ТО и Р АТ;
- **текущее планирование** предусматривает составление планов по решению конкретных задач авиакомпании.

В зависимости от предназначения и содержания планов используются следующие методы планирования.

Метод текстуального изложения плана – порядок выполнения планируемых мероприятий описывается текстуально в произвольной или специально разработанной форме (*Например, № п/п, мероприятия, кто проводит, место проведения, ответственные, отметка о выполнении, кто составил план*). Этот метод чаще всего используется при календарном и текущем планировании.

Графический - порядок выполнения планируемых мероприятий изображается в виде схем и графиков, иллюстрирующих распределение сил и средств и динамику выполнения работ.

Линейный - выполняемые работы изображаются в виде линий (лент), сведенных в таблицу, как правило, в масштабе времени. Этот метод используется при технологическом планировании. Линейные (*ленточные*) графики

очень просты, их несложно строить, и они обладают достаточной наглядностью. Но и есть недостатки:

- при большом количестве выполняемых работ (*60...70 разнотипных работ*) трудно проследить последовательность их выполнения и взаимосвязь между отдельными работами. Между тем основная сложность планирования технологического процесса заключается в необходимости отразить не количество работ, а их взаимосвязь;

- отсутствие замкнутой цепи на таких графиках затрудняет проведение математической обработки показателей всего комплекса работ.

Все это снижает возможности по отработке оптимизированных планов выполнения работ на АТ и эффективному управлению этими процессами.

Сетевой метод. В 50-х годах прошлого столетия было разработано два схожих метода управления работами по реализации проектов. «Проект – это совокупность задач или мероприятий, связанных с достижением запланированной цели, которая обычно имеет уникальный и не повторяющийся характер».

Метод критического пути («СРМ» - *Critical Path Method*). Был предложен фирмами «Дюпон» и «Ремингтон Рэнд Юнивак» для руководства выполнением крупных проектов по модернизации заводов фирмы «Дюпон». Важной предпосылкой применения метода СРМ является предположение о том, что время выполнения каждой работы точно известно.

Метод оценки и обзора программы («PERT» - *Program Evaluation Review Technique*). Был создан корпорацией «Локхид» в ВМС США при разработке ракетной системы «Поларис». Метод PERT ориентирован на анализ проектов, для которых продолжительность выполнения всех или некоторых работ не удается определить точно. Он применяется при проектировании и внедрении новых систем. В таких проектах многие работы не имеют аналогов. В результате возникает неопределенность в сроках выполнения проекта в целом.

В советской практике оба метода, а также их более сложные модификации объединяются понятием «методы сетевого планирования и управления». В методе критического пути и методе «PERT» проекты рассматриваются как сети отдельных событий и работ.

Работа в этих системах представляет собой любой элемент проекта, на выполнение которого требуется время и который может задержать выполнение других работ. Основное различие двух методов заключается в различном подходе к длительности операций. Метод критического пути исходит из того, что длительность операций можно оценить с достаточно высокой степенью точности и определенности.

В настоящее время более широко применяют метод критического пути, а не метод оценки и пересмотра планов. В последние годы произошел «настоящий взрыв в применении систем управления на основе сетевых графиков при реализации мелких и средних по масштабам проектов, возник большой спрос на руководителей проектов, понимающих принципы аналитических методов сетевого планирования и владеющих ими».

2.3.2 Сетевой метод планирования работы

Основу сетевого планирования составляет сетевой график, разработка которого базируется на теории ориентированных графов, теории вероятностей и математической статистике.

Сетевой график представляет собой графическое изображение комплекса работ, представленных в определенной последовательности их выполнения и взаимосвязи.

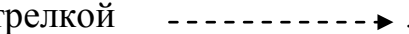
Основные элементы сетевого графика: «*работа*», «*событие*», «*путь*».

Понятие работа включает:

- **действительная работа** – любой трудовой процесс, сопровождающийся затратами труда, времени и материальных ресурсов. Изображается сплошной стрелкой



- **фиктивная работа** – логическая связь между выполняемыми работами, указывающая на невозможность или недопустимость выполнения какой-либо работы до окончания другой, т.е. фиктивная работа показывает, что возможность начала другой работы непосредственно зависит от окончания другой. Изображается пунктирной стрелкой

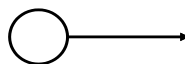


- **работа ожидания** – пассивный процесс, связанный только с затратами времени, но не требующий затрат труда и материальных ресурсов (*процесс высухания ЛКП, прогрев аппаратуры*). На безмасштабных графиках изображается сплошной стрелкой, на масштабных – пунктирной;

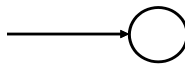
Каждая работа имеет начальное и конечное событие.

Событие – факт, фиксирующий начало или окончание работы, или нескольких работ (*самолет заправлен топливом*). Обозначение события в сетевом графике в виде геометрической фигуры (*обычно – в виде окружности*).

Начальным событием называется событие, за которым следует данная работа.



Конечным является событие, которому непосредственно предшествует данная работа.



Начальное событие, характеризующее начало выполнения всего комплекса работ, называется **исходным** событием. Его особенность – не имеет предшествующих работ.

Конечное событие, характеризующее окончание выполнения всего комплекса работ, называется **завершающим**.

В сетевом графике приняты обозначения:

i – текущий номер начального события (*стрелка выходит*);

j – текущий номер конечного события (*стрелка входит*).

При этом всегда $i < j$.

Путь - любая непрерывная технологическая последовательность работ между какими-то событиями. Различают:

L – полный путь. Путь называется полным, если маршрут берется от исходного события сетевого графика до завершающего. Продолжительность каждого полного пути определяется как сумма продолжительностей действительных работ.

Полный путь, имеющий максимальную продолжительность, называется критическим. На сетевом графике критический путь выделяется двойной линией, утолщенной или цветной.

Пути, имеющие продолжительности меньше критической, но близкие к ней, называются подкритическими.

На практике используют понятия: L_1 – **участки пути**, путь от исходного события до данного, L_2 – **участок пути**, путь от данного события до завершающего.

Сетевые графики могут строиться как безмасштабными, так и в масштабе времени. Пример сетевого графика из [1] представлен на рис 2.10.

Формирование исходных данных, порядок построения сетевого графика.

Построение сетевого графика на исследуемый процесс условно подразделяется на три основных этапа: выполнение подготовительных мероприятий (*операций*), составление непосредственно сетевого графика, расчет параметров сетевого графика и их анализ и оптимизация.

I этап начинается с уяснения поставленной задачи, определяющей конечную цель планируемой разработки, на достижение которой должны быть направлены усилия всех специалистов, и включает:

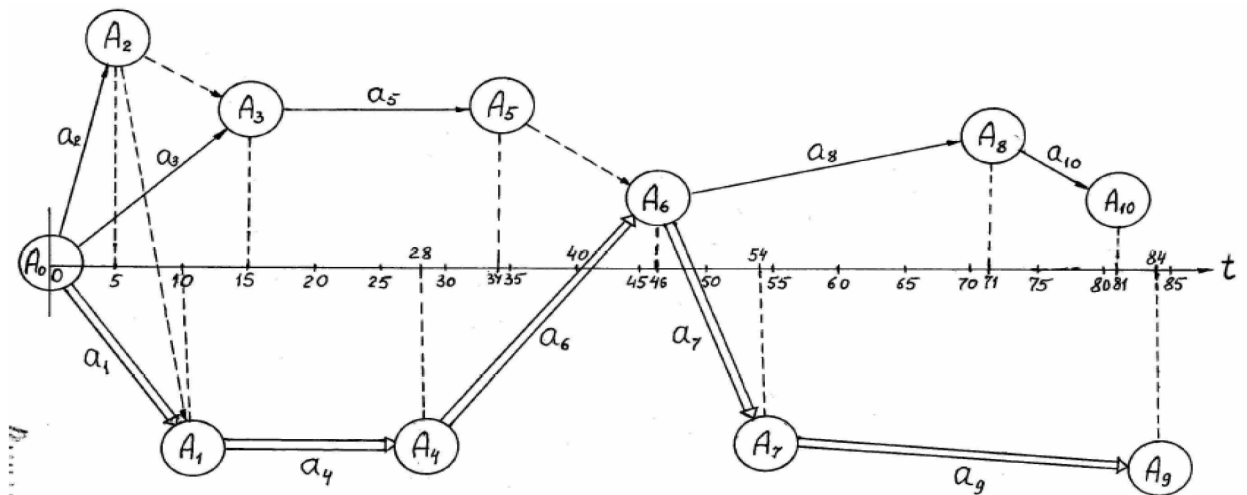


Рис. 2.10 Пример сетевого графика

1) составление перечня работ, которые должны быть выполнены для достижения поставленной цели;

2) определение временных параметров выполнения каждой работы (*продолжительность работ*) и потребных сил и средств.

Перечни работ и временные параметры этих работ сводятся в табл.2.1 исходных данных (*виртуальный пример*).

Таблица 2.1

Перечни работ и временные параметры работ

Код работы	Наименование работы	Продол-ность работы, мин	Шифр исп-ля/ число исп-лей
1	2	3	4
1.1	Подготовка КПА, инструмента	3	1.1
1.2	Осмотр самолета по маршруту	15	1.1
1.3	Проверка заправки (зарядки) систем самолета	10	1.1
1.4	Дозаправка самолета топливом	35	1.1/2
1.5	Заключительные работы	15	1.1
1.6	Оформление документации	10	1.1

Каждая работа должна иметь достоверную временную оценку ее выполнения. Следует помнить, что качество сетевого графика, эффективность планирования, а в дальнейшем и управления в значительной степени зависят от достоверности временных оценок запланированных работ. Для определения и уточнения среднего времени выполнения каждой запланированной работы необходимо использовать справочно-нормативные и статистические материалы. При необходимости производится дополнительное хронометрирование отдельных операций.

Каждой работе присваивается свой определенный шифр (*код*), который записывается в таблицу перечня работ и затем проставляется на стрелках сетевого графика.

Если расчет параметров сетевого графика предполагается проводить с учетом случайного характера времени выполнения работ, то в таблицу помещают для каждой работы не только среднее время выполнения работ t_{cp} , но и среднее квадратичное отклонение $\sigma_{t_{cp}}$.

II этап заключается в составлении общего сетевого графика в целом и включает:

- 1) составление частных сетевых графиков;
- 2) «сшивание» сетевого графика;
- 3) проверку правильности построения общего сетевого графика;
- 4) проведение единой нумерации событий общего сетевого графика.

Рассмотрим второй этап более подробно.

Составление частных сетевых графиков обычно не представляет сложности.

Основой для «сшивания», т.е. объединения, частных графиков являются: общие события и работы, логическая последовательность выполнения работ и их взаимосвязь. При этом должны учитываться возможные логические связи

между операциями, очередность их выполнения в соответствии с технологией процесса, допустимым фронтом работ, мерами безопасности и т.д.

Следует помнить, что выбор рационального характера связей между событиями и работами при составлении общего графика является одним из наиболее сложных этапов планирования. Конечная цель исследуемого процесса может быть достигнута при самых различных сочетаниях промежуточных событий и работ.

Искусство руководителя разработкой и заключается в том, чтобы из множества возможных вариантов выбрать наилучший, обеспечивающий выполнение всего плана в минимальное время и с наименьшими затратами сил и средств.

При относительно простом комплексе работ на самолете и наличии соответствующего опыта в составлении сетевых графиков возможно частные графики и не составлять.

После сшивания сетевого графика необходимо проверить правильность его составления. При этом необходимо учитывать следующие особенности:

1. Каждая действительная работа и работа ожидания на сетевом графике имеет обозначение:

$$\begin{array}{c} \underline{11} \longrightarrow \\ 10 \end{array}$$

- над стрелкой обозначается шифр работы
- под стрелкой - время ее выполнения.

2. На рис. 2.11 показаны возможные ошибки, допускаемые при составлении сетевых графиков.

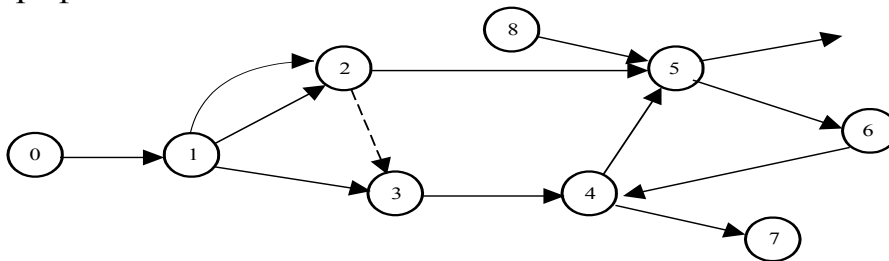


Рис. 2.11 Характерные ошибки сетевого графика

На графике не должно быть:

- «повисающих» (*тупиковых*) событий, т.е. таких событий (*кроме завершающего*), из которых не выходит ни одной работы (рис. 2.13 событие 7);
- событий (*кроме исходного*), в которые не входит ни одна работа (*событие 8*);
- замкнутых контуров (*циклов*), т.е. такого расположения работ и событий, при котором, следуя по направлению стрелки, возможен возврат к уже пройденному событию (*цикл 5- 6 - 4 - 5*);
- более одной работы между двумя последовательными событиями (*события 1-2*).

Если между двумя событиями (рис. 2.12) заключены две и более параллельные работы, то такие связи должны быть дополнены фиктивными работами.



Рис. 2.12 Возможное расположение событий

Расположение событий должно, по возможности, исключить взаимные пересечения работ, а направление стрелок должно быть от исходного до завершающего события (*т.е. стрелки имели бы направление слева направо*).

3. Процесс построения сетевого графика обычно заканчивается нумерацией событий (рис 2.13). При этом обычно используется метод ранжирования (*вычеркивания стрелок*).

Начальному (*исходному*) событию присваивается номер «0» и производится вычеркивание одной чертой всех стрелок, выходящих из этого события (рис.2.15). Его именуют событием нулевого ранга.

Все работы, выходящие из события нулевого ранга, именуют работами первого ранга. Те события, в которые входят только работы первого ранга, называются событиями первого ранга, им присваиваются номера 1, 2 и т.д. Нумерация событий первого ранга производится произвольно.

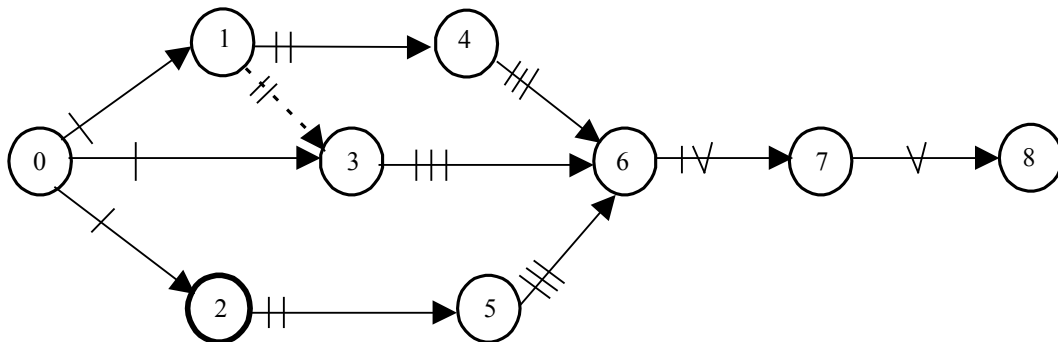


Рис. 2.13 Нумерация событий

Работы, выходящие из событий первого ранга, вычеркиваются двумя черточками. События, в которые входят только работы второго и предыдущего рангов, называются событиями второго ранга и снова нумеруются следующие события т.д. до завершающего.

III этап включает в себя расчет основных параметров общего сетевого графика, их анализ и оптимизацию по времени всего комплекса планируемых работ.

Расчет параметров сетевого графика.

Основными параметрами сетевого графика являются:

- 1) продолжительность пути, критический путь - $t(L)$, $t_{кр}(L_{кр})$;
- 2) ранний срок свершения события - $t_p(i)$,

- поздний срок свершения события - $t_{п}(i)$;
- 3) резерв времени свершения события – $R(i)$;
 - 4) полный резерв времени работы - $R_{п}(i, j)$;
 - 5) полный резерв времени пути - $R_{п}(L_i)$;
 - 6) коэффициент напряженности пути – $K_{н}(L_i)$.

В общем случае *продолжительность пути* равна сумме времен работ, лежащих на данном пути:

$$t(L_{o-k}) = \sum_{i=o}^k t(i, j),$$

где k - количество работ, лежащих на данном пути;

$t_{i,j}$ - продолжительность (время) выполнения работы, лежащей на данном пути;

$o-k$ - индексы исходного и завершающего событий.

От исходного события до завершающего может быть несколько путей.

Критический путь определяет минимальное время выполнения всего комплекса работ.

Наиболее простым (но трудоемким) способом определения $t_{кр}$ является метод перебора всех путей, но удобнее находить продолжительность критического пути посредством определения ранних сроков свершения всех событий.

Ранний срок свершения j -го события $t_p(j)$ - наибольший по продолжительности из путей, ведущий от исходного события к данному.

$$t_p(j) = t [L_1 \max(o-j)].$$

Следовательно, $t_p(j)$ – это минимальное время, необходимое для выполнения всех работ, предшествующих данному j -му событию, т.е. при заданной последовательности выполнения работ и их взаимосвязях получить меньшее время выполнения всех работ нельзя.

Зная ранний срок свершения i -х событий $t_p(i)$ и продолжительность работ $t(i,j)$, можно вычислить ранний срок свершения j -го события:

$$t_p(j) = \max [t_p(i) + t(i,j)].$$

Расчет ранних сроков свершения событий производится слева направо от исходного до завершающего события (рис. 2.14). Ранний срок свершения завершающего события и определяет продолжительность критического пути.

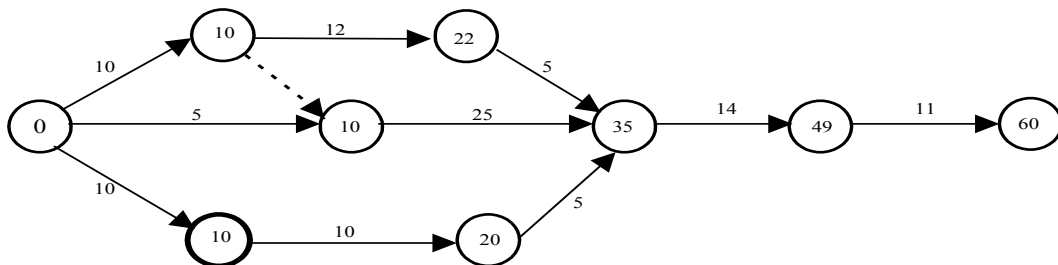


Рис. 2.14 Расчет ранних сроков свершения событий

Поздний срок свершения i -го события $t_{п}(i)$ - есть разность между продолжительностью критического пути $t_{кр}$ и наибольшим по продолжительности из путей, ведущих от данного события к завершающему.

$$t_{п}(i) = t_{кр} - t [L_2 \max(i,k)],$$

где k - индекс завершающего события.

Следовательно – $t_{п}(i)$ - максимально допустимое время, превышение которого вызовет задержку выполнения всего комплекса работ, т.е.

$$t_{п}(i) = \min [t_{п}(j) - t(i,j)].$$

Расчет $t_{п}(i)$ удобно вести справа налево (рис.2.15).

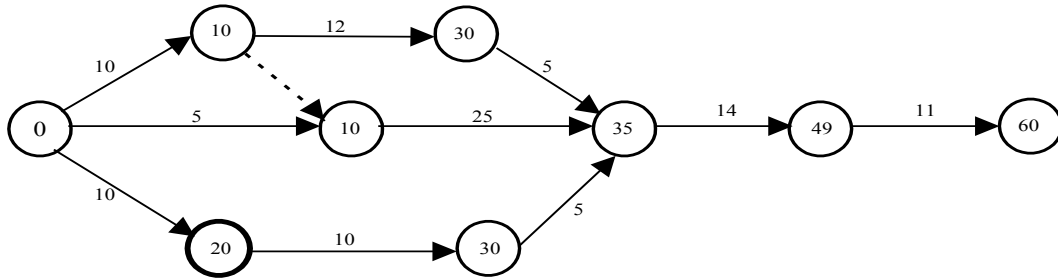


Рис. 2.15 Расчет поздних сроков свершения событий

Резерв времени события $R(i)$ - время, на которое может быть задержано свершение данного события без увеличения срока выполнения всего комплекса работ.

$$R(i) = t_{п}(i) - t_{р}(i),$$

т.е. $R(i)$ - есть разность между поздним и ранним сроком свершения события. Для всех событий, лежащих на критическом пути:

$$t_{п}(i) = t_{р}(i), \quad \text{т.е. } R(i) = 0.$$

Следовательно, путь, у которого резервы равны нулю, является критическим.

Полный резерв времени работы $R_{п}(i, j)$ - время, на которое может быть увеличено время выполнения данной работы, при котором длина критического пути остается неизменной (рис.2.16):

$$R_{п}(i,j) = t_{п}(j) - t_{р}(i) - t(i,j),$$

т.е. работа (i, j) не может быть начата раньше, чем свершится событие (i) , а конец работы (i, j) - не позже, чем $t_{п}(j)$.

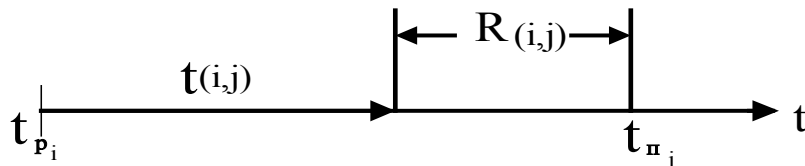


Рис. 2.16 Расчет полного резерва времени работы

Полный резерв времени пути $R_{\Pi}(L_i)$ - разность между длиной критического и данного пути:

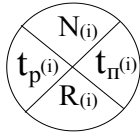
$$R_{\Pi}(L_i) = t_{кр} - t(L_i).$$

Коэффициент напряженности $K_H(L_i)$ - есть отношение длины данного пути к длине критического пути:

$$K_H(L_i) = \frac{t(L_i)}{t_{кр}}.$$

Весьма наглядным и удобным методом для расчета параметров сетевого графика является четырехсекторный метод. Метод предусматривает выполнение расчетов непосредственно на графике.

Каждая окружность, изображающая событие на сетевом графике, разделяется на четыре сектора, в которых записываются значения параметров $t_p(i)$, $t_{\Pi}(i)$, $R(i)$, $N(i)$.



где $N(i)$ – номер события,

и выполняется расчет в следующей последовательности (рис.2.17):

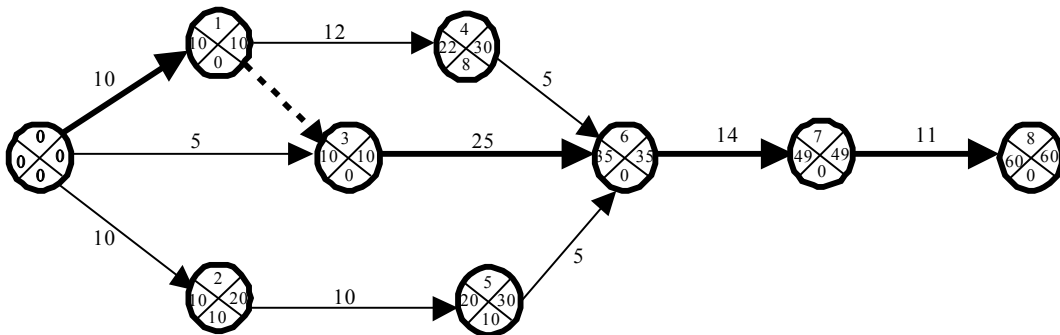


Рис. 2.17 Расчет параметров сетевого графика

- 1) производится единая нумерация событий;
- 2) рассчитываются значения t_p всех событий;
- 3) определяется длина критического пути, равная значению t_p завершающего события;
- 4) определяется значение t_{Π} для всех событий;
- 5) выделяются события, лежащие на критическом пути, т.е. события, для которых $t_p = t_{\Pi}$ и определяется критический путь;
- 6) рассчитывается резерв времени для каждого события.

2.3.3 Оптимизация сетевых графиков

При планировании работы ИАС с использованием сетевых графиков обычно предъявляется требование, чтобы планируемый комплекс работ мог быть выполнен за заданное время и была обеспечена равномерная и достаточно полная загрузка исполнителей и технических средств. Однако составить сразу сетевой график, удовлетворяющий этим требованиям, как правило, затруднено.

Поэтому после первоначального составления и расчета параметров сетевого графика производится его дальнейшее совершенствование, т.е. – оптимизация.

Оптимизация сетевого графика - выбор варианта сетевого графика, который обеспечивает выполнение комплекса работ за заданное или минимальное время при сохранении равномерной и достаточно полной загрузки исполнителей и технических средств.

Основные задачи оптимизации:

- 1) изыскание возможности сокращения общей продолжительности выполнения всего комплекса работ (*т.е. уменьшения $t_{кр}$*);
- 2) обеспечение равномерной напряженности путей ($K_n(L_i)$);
- 3) обеспечение полноты и равномерной загрузки специалистов и СНО (*коэффициент загрузки личного состава K_z , загрузки СНО $K_{z\ сно}$*).

Приемы оптимизации: сдвиг работ графика за счет резервов времени, перераспределение имеющихся сил и средств, ввод дополнительных ресурсов.

При составлении сетевого графика необходимо обеспечить условие, при котором $t_{кр}$ было бы не более заданного (директивного). Если имеются данные о времени выполнения работ как о случайных величинах, то целесообразно определить значения вероятности выполнения комплекса работ за заданное время.

Для эффективного применения метода сетевого планирования необходимо соблюдать два условия:

- 1) основные элементы техники сетевого планирования должны быть изучены в совершенстве;
- 2) сетевой метод может быть введен в практику только после тщательной подготовки личного состава, при полном понимании целей управления и сложности планируемых процессов.

Жизнь подтвердила, что поверхностное отношение к методу сетевого планирования неминуемо ведет к дискредитации метода и провалу работ.

Лекция №2.4

Раздел 2. Анализ систем методами исследования операций

Тема № 2.4 Методы математического программирования

2.4.1. Математическое программирование как инструмент решения задач выбора.

2.4.2 Линейное программирование.

2.4.2.1 Постановка задачи линейного программирования.

2.4.2.2. Геометрическая интерпретация линейного программирования.

2.4.3. Нелинейное и динамическое программирование.

2.4.3.1. Задачи нелинейного программирования.

2.4.3.2. Динамическое программирование.

Заключение по курсу. Рекомендации по использованию материалов дисциплины в специальных курсах.

2.4.1 Математическое программирование как инструмент решения задач выбора

В качестве математических средств принятия решений используются:

- перебор вариантов решения для выбора из них оптимального;
- математическое программирование;
- известные в теории специальные модели (*очереди, игр и др.*).

Чтобы воспользоваться математическим аппаратом, необходимо сначала сформулировать интересующую нас проблему как математическую задачу, придав количественные оценки возможным вариантам и придав количественный смысл словам «лучше», «хуже».

Математическое программирование является инструментом решения разнообразных задач выбора. Математическое программирование в отличие от классических методов решения экстремальных задач (*их называют задачи оптимизации*) основное внимание уделяет вопросам ограничений на область изменения переменных. Само наименование «математическое программирование» связано с тем, что целью решения задач является выбор *программы* действий.

Термин «программирование» в названии темы ничего общего с термином «программирование (*т.е. составление программ*) для ЭВМ» не имеет, так как в общем случае такая дисциплина как «линейное программирование» возникла еще до того времени, когда ЭВМ стали широко применяться при решении математических, инженерных, экономических и др. задач.

Термин «линейное программирование» возник в результате неточного перевода английского «*linear programming*». Одно из значений слова «*programming*» - составление планов, планирование.

Следовательно, правильным переводом «linear programming» было бы не «линейное программирование», а «линейное планирование», что более точно отражает содержание дисциплины.

Однако, термин линейное программирование, нелинейное программирование и т.д. в нашей литературе стали общепринятыми.

Математическое программирование («планирование») – это раздел математики, занимающийся разработкой методов отыскания экстремальных значений функции, на аргументы которой наложены ограничения.

Методы математического программирования используются в экономических, организационных, военных и др. системах для решения так называемых распределительных задач. Распределительные задачи возникают в случае, когда имеющихся в наличии ресурсов не хватает для выполнения каждой из намеченных работ эффективным образом и необходимо наилучшим образом распределить ресурсы по работам в соответствии с выбранным критерием оптимальности.

В общей постановке задачи математического программирования требуется найти неотрицательные значения n переменных x_1, x_2, \dots, x_n , которые удовлетворяют m уравнениям или неравенствам

$$a_{ij}(x_1, x_2, \dots, x_n) \{ \leq, =, \geq \} b_i \\ i = 1, 2, \dots, m, j = 1, 2, \dots, n$$

и максимизируют или минимизируют целевую функцию

$$F(x) = f(x_1, x_2, \dots, x_n) = \{ \max \vee \min \}$$

Таким образом, формулировка задачи математического программирования состоит в следующем.

Вводится скалярная функция векторного аргумента $F(x)$. Эту функцию в зависимости от конкретного вида решаемой задачи **называют целевой функцией, критерием эффективности** или **показателем качества**.

Постановка задачи и методы исследования существенно зависят от свойств целевой функции.

Целевая функция позволяет выбирать наилучший вариант - из множества возможных. Наилучший вариант доставляет целевой функции экстремальное значение. Это может быть прибыль, объем выпуска или реализации, затраты производства, издержки обращения, уровень обслуживания или дефицитности, число комплектов, отходы и т. д.;

Эти условия следуют из ограниченности ресурсов, которыми располагает общество в любой момент времени, из необходимости удовлетворения насущных потребностей, из условий производственных и технологических процессов. Ограниченными являются не только материальные, финансовые и трудовые ре-

сурсы. Таковыми могут быть возможности технического, технологического и вообще научного потенциала.

В первом разделе, рассматривая вопрос «Задачи оптимизации», мы говорили. Наиболее просто, с математической точки зрения, случаи, когда целевая функция задается явной формулой и является при этом дифференцируемой функцией. В этом случае для поиска точек локального экстремума может быть использована производная.

Но в последние годы круг задач оптимизации, поставленных практикой, резко расширился. Во многих из них целевая функция не задается формулой, ее значения могут быть получены в результате сложных расчетов, братья из эксперимента и т.д.

Такие задачи являются более сложными, потому что для них нельзя провести исследование целевой функции с помощью производной. Приходится уточнять их математическую постановку и разрабатывать специальные методы решения, рассчитанные на широкое применение ЭВМ. Нередко потребности превышают возможности их удовлетворения. Математические ограничения выражаются в виде уравнений и неравенств. Их совокупность образует **область допустимых решений** (*область экономических возможностей*).

План, удовлетворяющий системе ограничений задачи, называется *допустимым*. Допустимый план, доставляющий функции цели экстремальное значение, называется *оптимальным*. Оптимальное решение, вообще говоря, не обязательно единственно, возможны случаи, когда оно не существует, имеется конечное или бесчисленное множество оптимальных решений.

Если целевая функция $F(x)$ и функции a_{ij} и b_i линейны, то соответствующая задача является задачей *линейного программирования*. Если хотя бы одна из упомянутых функций является нелинейной, то такая задача относится к задаче *нелинейного программирования*. Если в целевой функции $F(x)$ и функциях a_{ij} и b_i , определяющих ограничения, содержатся параметры с элементами неопределенности, то такая задача относится к задаче *стохастического программирования*. Задачи, решения которых достигаются в результате многих этапов, относятся к задачам *динамического программирования*.

2.4.2 Линейное программирование

Родоначальником ЛП считают академика Л.В. Канторовича, который в 30-е гг. XX в. предложил метод оптимизации решения экономических задач (*в частности, задачи раскроя фанеры*).

Термин «линейное программирование» был предложен Джорджем Бернардом Данцигом (1914-2005) в 1949г. для изучения теоретических и алгоритмических задач, связанных с оптимизацией линейных функций при линейных ограничениях (*выдающийся математик США, разработал симплексный метод решения ЗЛП. Наряду с советским математиком Л.В. Канторович, считается «отцом линейного программирования»*).

2.4.2.1 Постановка задачи линейного программирования.

В общем виде ЗЛП записывается следующим образом:

1) формулируется целевая функция (максимум или минимум);

$$F = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \rightarrow \max (\min)$$

Ранее эту функцию в зависимости от конкретного вида решаемой задачи мы называли целевой функцией, критерием эффективности или показателем качества. Решение одномерной задачи не представляет ни малейших затруднений. Здесь можно воспользоваться аппаратом производной - при конкретном значении x найти производную максимизируемой функции, приравнять ее нулю, решить полученное уравнение и вычислить значения функции в критической точке.

2) учитывается система ограничений в форме линейных уравнений и неравенств;

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \{ \leq, \geq, = \} b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \{ \leq, \geq, = \} b_2, \\ \dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \{ \leq, \geq, = \} b_m. \end{cases}$$

Коэффициенты

$a_{ij}, b_i, c_j, i = 1, 2, \dots, m, j = 1, 2, \dots, n$ - любые действительные числа (возможно 0).

3) учитывается требование неотрицательности переменных.

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$$

В случае, когда число переменных n равно числу ограничивающих уравнений m , система уравнений имеет единственное решение, которое может быть найдено с помощью матричной алгебры.

При $m < n$, совокупность значений $x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$, удовлетворяющих системе ограничений, образуют многомерную область *допустимых* решений (ОДР).

Значения же $\max (\min) F (x_1, x_2, \dots, x_n)$ обеспечиваются при определенных условиях и соответствует *оптимальному* решению или решениям (если они существуют).

Таким образом, решение ЗЛП состоит в поиске неотрицательных значений переменных, удовлетворяющих ограничениям, заданным системой линейных уравнений и неравенств, и отвечающих наибольшему (наименьшему) значению линейной функции переменных, рассматриваемой как целевой.

Рассмотрим простейший пример задачи ЛП.

Для производства изделий (*стульев и столов*) расходуется древесина:

- на один стул – 0.04 м^3 - a_{11} ;

- на один стол – 0.5 м^3 - a_{12} .

Изделия реализуются по отпускным ценам:

- за один стул – 0.4 тыс. руб. – C_1 ;

- за один стол – 2.8 тыс. руб. – C_2

Ресурсы мебельного предприятия ограничены 200 м^3 древесины.

Необходимо найти такую оптимальную производственную программу, чтобы выручка от реализации стульев и столов была максимальной при условии выполнения уже имеющегося заказа на изготовление 80 столов.

Экономико-математическая модель строится с учетом указанных расходов древесины, ограничений и условий и записывается следующим образом:

$$F(x_1, x_2) = C_1 x_1 + C_2 x_2 \rightarrow \max$$

$$x_1 \geq 0; x_2 \geq 80;$$

$$a_{11} x_1 + a_{12} x_2 \leq 200$$

- x_1 - оптимальное целое количество изготавливаемых стульев, $x_1 \geq 0$;

- x_2 - оптимальное целое количество изготавливаемых столов, $x_2 \geq 80$;

- ограничения по древесине, $0.04 x_1 + 0.5 x_2 \leq 200$;

- целевая функция

$$F(x_1, x_2) = 0.4 x_1 + 2.8 x_2 \rightarrow \max$$

Решение задачи: $x_1 = 4000$; $x_2 = 80$ (*графическим или другим методом*).

Проверим: $0.4 * 4000 + 2.8 * 80 = 1824$

$$0,04 * 4000 + 0,5 * 80 = 200 \quad \text{т.е то что надо } \leq 200$$

Наиболее важными формами задачи линейного программирования являются **каноническая и стандартная**.

В *канонической форме* задача является задачей на максимум (*минимум*) некоторой линейной функции **F**, ее система ограничений состоит *только из равенств (уравнений)*.

При этом переменные задачи x_1, x_2, \dots, x_n являются неотрицательными:

К канонической форме можно преобразовать любую задачу линейного программирования.

В стандартной форме задача линейного программирования является задачей на максимум (минимум) линейной целевой функции.

Система ограничений ее состоит из *одних* линейных неравенств типа « \leq » или « \geq ». Все переменные задачи неотрицательны.

Всякую задачу линейного программирования можно сформулировать в стандартной форме.

Всякое равенство в системе ограничений равносильно системе взаимоположных неравенств:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i, \\ -a_{i1}x_1 - a_{i2}x_2 - \dots - a_{in}x_n \leq -b_i, \end{cases}$$

Существует и другие способы преобразования системы равенств в систему неравенств, т.е. всякую задачу линейного программирования можно сформулировать в стандартной форме.

2.4.2.2 Геометрическая интерпретация линейного программирования

Геометрическая интерпретация подобных задач дает возможность наглядно представить их структуру, выявить особенности и открывает пути исследования более сложных свойств.

ЗЛП с двумя переменными всегда можно решить графически.

Однако уже в трехмерном пространстве такое решение усложняется, а в пространствах, размерность которых больше трех, графическое решение, вообще говоря, невозможно.

Следует заметить, что случай двух переменных не имеет особого практического значения, однако его рассмотрение проясняет свойства ЗЛП, приводит к идее ее решения, делает геометрически наглядными способы решения и пути их практической реализации.

Пусть дана задача

$$\max Z = c_1x_1 + c_2x_2 \quad (*)$$

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 \leq b_1 \\ a_{21}x_1 + a_{22}x_2 \leq b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 \leq b_m \end{array} \right\} \quad (**)$$

$$x_1 \geq 0, x_2 \geq 0 \quad (***)$$

Дадим геометрическую интерпретацию элементов этой задачи.

Пусть область допустимых решений ЗЛП (рис.2.18) - непустое множество, например, многоугольник $A_1 A_2 A_3 A_4 A_5 A_6$.

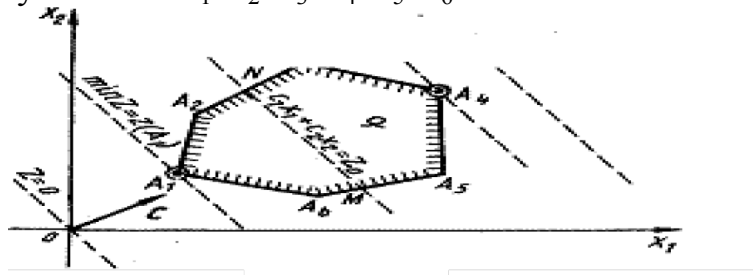


Рис. 2.18 Область допустимых решений ЗЛП

Выберем произвольное значение целевой функции $Z=Z_0$.

Получим $c_1x_1 + c_2x_2 = Z_0$.

Это уравнение прямой линии. В точках прямой NM целевая функция сохраняет одно и то же постоянное значение Z_0 . Считая в равенстве (*) Z параметром, получим уравнение семейства параллельных прямых, называемых линиями уровня целевой функции (*линиями постоянного значения*).

При нахождении решения могут встретиться случаи, изображенные на рис. 1- 4. Рис. 2.19 характеризует такой случай, когда целевая функция принимает максимальное значение в единственной точке A . Из рис. 2.20 видно, что максимальное значение целевая функция принимает в любой точке отрезка AB .

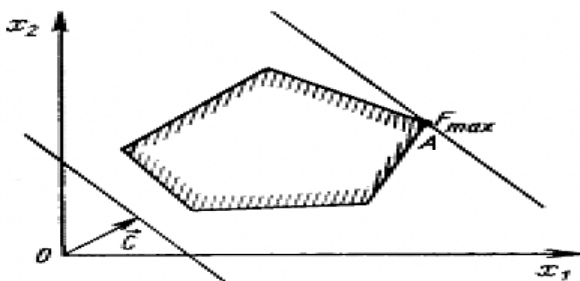


Рис. 2.19

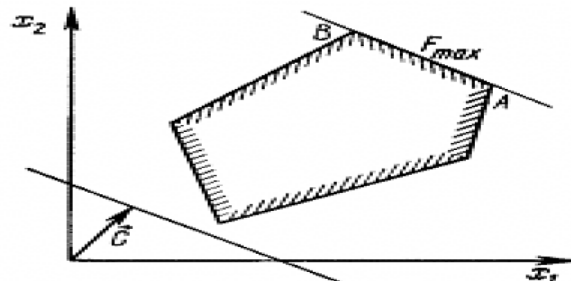


Рис. 2.20

На рис. 2.21 изображен случай, когда целевая функция не ограничена сверху на множестве допустимых решений, а на рис. 2.22 – случай, когда система ограничений задачи несовместна.

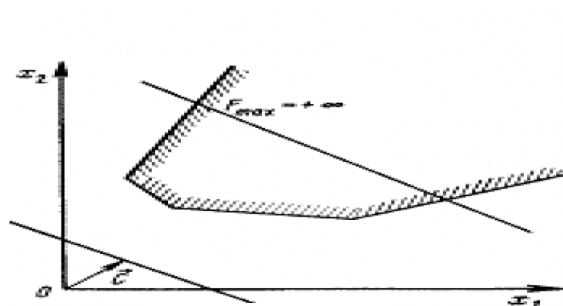


Рис. 2.21

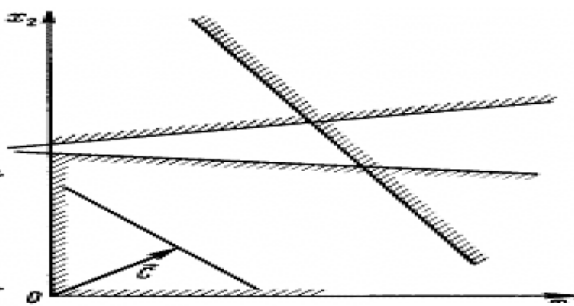


Рис. 2.22

2.4.3 Нелинейное и динамическое программирование

2.4.3.1 Задачи нелинейного программирования

Задачами нелинейного программирования называются задачи математического программирования, в которых нелинейны и (или) целевая функция, и (или) ограничения в виде неравенств или равенств.

Общих способов решения, для нелинейного программирования не существует. В каждом конкретном случае способ выбирается в зависимости от вида функции $F(x)$. Многие задачи нелинейного программирования могут быть приближены к задачам линейного программирования, и найдено близкое к оптимальному решению. В целом задачи нелинейного программирования относятся к трудным вычислительным задачам. При их решении часто приходится прибегать к приближенным методам оптимизации.

Перечислим свойства ЗНП, которые существенно усложняют процесс их решения по сравнению с задачами линейного программирования:

1. Множество допустимых решений может иметь очень сложную структуру (*например, быть невыпуклым или несвязным*).

2. Глобальный максимум (минимум) может достигаться как внутри множества допустимых решений, так и на его границах (*где он, вообще говоря, будет не совпадать ни с одним из локальных экстремумов*).

3. Целевая функция $F(x)$ может быть недифференцируемой, что затрудняет применение классических методов математического анализа.

В силу названных факторов задачи нелинейного программирования настолько разнообразны, что для них не существует общего метода решения.

Мощным средством для решения задач нелинейного программирования являются численные методы. Они позволяют найти решение задачи с заданной степенью точности.

2.4.3.2 Динамическое программирование

Динамическое программирование (*динамическое планирование*) представляет собой математический метод оптимизации решений, специально приспособленный к многошаговым операциям (*процессам, характеризующим систему*).

Процесс, исследуемый методом динамического программирования, может явно зависеть от времени и естественно распадаться на ряд шагов или этапов.

Примером такой деятельности могут быть разнообразные задачи распределения ресурсов в процессе планирования хозяйственной деятельности группы предприятий. При этом естественным шагом является хозяйственный год.

В самых различных областях деятельности приходится иметь дело с необходимостью поэтапного принятия решений для достижения некоторой конечной цели. Если для одношаговых процессов принимаемые решения, как

правило, относительно просто, то в многошаговых процессах структура решения несравнимо сложнее.

Так, если в лесном хозяйстве руководствоваться критерием максимального получения максимальной прибыли в текущем году без учета перспективы, то оптимальное решение сведется к требованию «рубь побольше».

Сегодня у цивилизованных людей не возникает сомнений в необходимости использования для управления сложными системами методов математического моделирования, которое подчас дает решения, неожиданные не только для начинающего, но и для опытного руководителя.

Очевидно, что никакая математическая модель не может описать действительность с исчерпывающей полнотой (*как утверждал небезызвестный Козьма Прутков, «никто не может объять необъятное»*). Более того, рост сложности модели требует более сложного математического аппарата ее анализа.

Потому, согласно Р. Беллману, «Ученый должен идти прямой и узкой тропой между Западными Переупрощения и Болотом Переусложнения».

Последствия переупрощения очевидны, учет же большого количества факторов приводит к известному «проклятию размерности», делающему задачу выбора оптимальной политики практически неразрешимой.

Пусть планируется деятельность группы предприятий на N лет. Здесь шагом является один год.

В начале 1-го года на развитие предприятий выделяются средства, которые должны быть как-то распределены между этими предприятиями. В процессе их функционирования выделенные средства частично расходуются. Каждое предприятие за год приносит некоторый доход, зависящий от вложенных средств. В начале года имеющиеся средства могут перераспределяться между предприятиями: каждому из них выделяется какая-то доля средств.

Ставится вопрос: как в начале каждого года распределять имеющиеся средства между предприятиями, чтобы суммарный доход от всех предприятий за N лет был максимальным?

Перед нами типичная задача динамического программирования, в которой рассматривается управляемый процесс функционирования группы предприятий. Управление процессом состоит в распределении (*и перераспределении*) средств. Управляющим воздействием (УВ) является выделение каких-то средств каждому из предприятий в начале года.

УВ на каждом шаге должно выбираться с учетом всех его последствий в будущем. УВ должно быть дальновидным, с учетом перспективы. Нет смысла выбирать на рассматриваемом шаге наилучшее УВ, если в дальнейшем это помешает получить наилучшие результаты других шагов. УВ на каждом шаге надо выбирать «с заглядыванием в будущее», иначе возможны серьезные ошибки.

В других случаях идущий во времени процесс является непрерывным (*например, движение самолетов или ракеты*). В этом случае этот процесс условно разбивают на этапы, каждый из которых занимает какой-то отрезок Δt .

Наконец, метод динамического программирования может быть применен для решения задач системного анализа, в частности, формирования структуры системы. В этом случае разбиение на шаги производится не по времени, а по любому другому признаку, например, по элементам системы.

В основе метода ДП лежит специфический принцип оптимальности, определяющий стратегию поиска оптимального управления.

Метод сформулирован в 1957 году выдающимся американским математиком Р. Беллманом в книге «Динамическое программирование» и определяет фундаментальное свойство оптимальной стратегии (*политики, поведения*).

Принцип формулируется следующим образом: оптимальное управление не зависит от предыстории процесса изменения состояния системы, а определяется лишь ее состоянием в рассматриваемый момент времени. Этот принцип определяет требования, предъявляемые к последующему движению системы, начиная с текущего момента времени и до конца процесса управления.

Говоря по-другому - «оптимальное поведение обладает тем свойством, что, каковы бы ни были первоначальное состояние и решение в начальный момент, последующие решения должны составлять оптимальное поведение относительно состояния, полученного в результате начального решения».

Это определение можно интерпретировать самым примитивным высказыванием: если вы намерены добиться наилучшего эффекта вашей многолетней деятельности, то на любом ее этапе, независимо от того, в каких состояниях вы оказывались и какие действия предпринимали, действуйте оптимально, насколько вам позволяет то состояние, в которое вы загнали себя предыдущими действиями.

По своему принципу динамическое программирование напоминает метод «разделяй и властвуй». Задача делится на подзадачи, которые либо являются очевидными, либо сводятся к своим подзадачам и т.д. Еще в XVII веке Р. Декарт писал: «Расчлените каждую изучаемую вами задачу на столько частей (...), сколько потребуется, чтобы их было легко решить».

И еще можно добавить - если на любом этапе своей целенаправленной деятельности руководствоваться стремлением к достижению некоторой высшей цели, то вся последовательность действий будет оптимальна.

Таким образом, принцип оптимальности определяет особенности политики, направленной не на получение сиюминутной выгоды, а на достижение некоторой удаленной цели, т.е. использование этого принципа гарантирует, что решение, выбранное на любом шаге, является не локально лучшим, а лучшим с точки зрения задачи в целом.

По существу, принцип оптимальности определяет порядок поэтапного решения, допускающий декомпозицию задачи с помощью рекуррентных соотношений. Рекуррентными соотношениями называют формулы, выражающие очередной член последовательности через один или несколько предыдущих членов.

Метод ДП решает общую оптимизационную задачу путем поиска оптимального управления на каждом шаге процесса. Иначе говоря, общая оптими-

зационная задача в методе ДП разбивается на ряд малых оптимизационных задач. Метод динамического программирования позволяет свести N -мерную задачу оптимизации к совокупности задач меньшей размерности (*очевидно, что легче решить 20 одномерных задач оптимизации, чем одну 20-мерную*).

В [1] излагается общий подход к проблеме управления системой методом динамического программирования.

Мы рассмотрим применение метода ДП на примере.

Рассмотрим известную *«задачу о загрузке корабля»*, которую в других эквивалентных постановках называют «задачей о рюкзаке» или «задачей линейного раскроя».

Пусть имеется корабль грузоподъемности W , который может быть загружен предметами N типов.

Если обозначить через W_k и C_k вес и ценность предмета k -го типа и через X_k количество таких предметов, то можно поставить задачу загрузки корабля грузом максимальной ценности в виде:

$$\begin{aligned} \text{максимизировать} \quad & \sum_{k=1}^N C_k X_k \\ \text{при условиях} \quad & \sum_{k=1}^N W_k X_k \leq W \quad X_k \geq 0 \quad \text{при всех } k \end{aligned}$$

Поскольку нам безразлично, в каком порядке грузить наши предметы, то вообразим искусственный n -шаговый процесс, на первом шаге которого загружается X предметов n -го типа (*на втором $n-1$ -го, затем $n-2$ -го и на последнем - первого типа*).

Тогда, обозначив через $F_n(W)$ суммарную ценность загрузки в n -шаговом процессе при заданной грузоподъемности W и при использовании оптимальной политики, *на основе принципа оптимальности* сводим задачу к системе рекуррентных соотношений

$$F_n(W) = \max[C_n X + F_{n-1}(W - W_n X)] \quad n = 2, 3, \dots, N:$$

$$F_1(W) = \max[C_1 X] \quad .$$

где область максимизации определяется целыми значениями X в диапазоне от нуля до целой части отношения W / W_n .

Вычисление оптимального процесса на основе уравнений Беллмана в каждом конкретном случае может оказаться непростой задачей, требующей навыка и изобретательности.

ЛИТЕРАТУРА

1. Кабков П.К. «Исследование операций и системный анализ». Учебное пособие – М.: МГТУ ГА, 2005.- 85с.
2. Демидов Ю.М., Колядов Д.В. Исследование операций. Учебное пособие. – М.: МГТУ ГА, 2013.
3. Волкова В.Н., Денисов А.А. Теория систем и системный анализ. Учебник для бакалавров. – М.: Москва, «Юрайт», 2012. – 860с.
4. Вентцель Е.С. Теория вероятностей. Учебник. – М.: Москва, «Изд. КноРус», 2010. – 654с.
5. Вентцель Е.С. «Исследование операций: задачи, принципы, методология». – М.: Москва, «Изд. КноРус», 2010. – 192с
6. Романов В.Н. Системный анализ для инженеров. Учебное пособие. - М.: Изд. СПб: СЗГЗТУ, 2006. - 186с.
7. Партыка Т.Л., Попов И.И. Математические методы. Учебник. – М.: Москва, «Форум-Инфра-М»,2007. – 464с.
8. Голубков Е.П. Системный анализ как методологическая основа принятия решений. Менеджмент в России и за рубежом №3, 2003.
9. М. Мескон, М.Альберт, Ф. Хеудори. Основы менеджмента. Перевод с английского. – М.: Москва, 2005.- «Дело». – 718с.
10. Волкова В.Н., Емельянова А.А. Теория систем и системный анализ в управлении организациями. Справочник. – М.: Москва, «Финансы и статистика»,2006. – 845с.
11. Вентцель Е.С., Овчаров Л.А. Задачи и упражнения по теории вероятностей. М.: Москва, 2003.- «ACADEMA». – 443с.

СОДЕРЖАНИЕ

Введение	3
I раздел. Теория систем и системный анализ	4
1.1 Основы системного анализа	4
1.1.1 Из истории исследования операций и системного анализа	4
1.1.2 Исследование операций как методологическая основа принятия решений	6
1.1.3 Системный анализ как последовательная реализация системного подхода	8
1.1.4 Методология системного подхода	9
1.2. Основы теории систем, основные понятия и определения	11
1.2.1 Понятие «система»	11
1.2.2 Классификация систем	12
1.2.3 Структура системы. Виды структур	13
1.2.4 Примеры систем воздушного транспорта и их краткая характеристика	16
1.3 Эффективность систем и показатели их качества	18
1.3.1 Понятия «эффективность» и «качество» систем	18
1.3.2 Подходы к выбору показателей и критериев эффективности и качества	21
1.4 Экстремальные задачи системного анализа	25
1.4.1 Научный подход к принятию решения	25
1.4.2 Задачи выбора и задачи оптимизации	28
1.4.3 Однокритериальные и многокритериальные задачи оптимизации	29
1.4.4 Множество Парето	30
Раздел 2. Анализ систем методами исследования операций	32
2.1 Анализ систем по схеме марковских случайных процессов	32
2.1.1. Марковские случайные процессы	32
2.1.2. Марковская цепь	34
2.1.3. Марковские процессы с дискретным состоянием и непрерывным временем	35
2.1.4. Полумарковские процессы	38
2.2 Анализ систем массового обслуживания	40
2.2.1. Компоненты систем массового обслуживания	40
2.2.2. Классификация систем массового обслуживания	42
2.2.3. Показатели качества обслуживания СМО	43
2.2.4. Анализ систем массового обслуживания с отказами	45
2.2.4.1 Система М/М/1/0 – одноканальная СМО с отказами	45
2.2.4.2 Система М/М/n/0 – многоканальная СМО с отказами	46
2.2.5. Анализ систем массового обслуживания с ожиданием	48
2.2.5.1 Системы М/М/1/m и М/М/1/∞ – одноканальные СМО с ожиданием	48

2.2.5.2 Системы $M/M/n/m$ и $M/M/n/\infty$ – многоканальные СМО с ожиданием	52
2.3 Метод сетевого планирования	55
2.3.1 Классификация методов планирования работы	55
2.3.2 Сетевой метод планирования работы	57
2.3.3 Оптимизация сетевых графиков	65
2.4 Методы математического программирования	66
2.4.1. Математическое программирование как инструмент решения задач выбора	66
2.4.2 Линейное программирование	68
2.4.2.1 Постановка задачи линейного программирования	69
2.4.2.2. Геометрическая интерпретация линейного программирования	71
2.4.3. Нелинейное и динамическое программирование	73
2.4.3.1. Задачи нелинейного программирования	73
2.4.3.2. Динамическое программирование	73
Литература.....	77

ЯБЛОНСКИЙ Сергей Николаевич,
ДАЛЕЦКИЙ Станислав Владимирович, МАТЮХИН Константин Николаевич

ИССЛЕДОВАНИЕ ОПЕРАЦИИ И СИСТЕМНЫЙ АНАЛИЗ
Тексты лекций

Подписано в печать 17.03.2017 г.

Печать офсетная
4,65 усл.печ.л.

Формат 60x84/16
Заказ № 1725/157

3,81 уч.-изд. л.
Тираж 30 экз.

*Московский государственный технический университет ГА
125993 Москва, Кронштадтский бульвар, д. 20
ООО «ИПП «ИНСОФТ»
107140, г. Москва, 3-й Красносельский переулок д.21, стр. 1*