

Министерство сельского хозяйства Российской Федерации

**Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Саратовский государственный аграрный университет им. Н.И. Вавилова»**

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ

краткий курс лекций

для обучающихся IV курса

**Направление подготовки
35.03.07 Технология производства и
переработки сельскохозяйственной продукции**

**Профиль подготовки
Технологии пищевых производств в АПК**

Саратов 2017

УДК 004
ББК 32.97
Р94

Рецензенты:

Доцент кафедры «Электроника колебаний и волн», кандидат физико-математических наук,
НИУ СГУ им. Н.Г. Чернышевского.

А.В. Толстиков

Заведующая кафедрой «Математика, математическое моделирование», кандидат физико-математических наук, доцент, ФГБОУ ВО «Саратовский ГАУ» им. Н.И. Вавилова

Г. Н. Камышова

Р94

Математическое моделирование технологических процессов: краткий курс лекций для обучающихся направления подготовки 35.03.07 Технология производства и переработки сельскохозяйственной продукции / Сост.: А.В. Розанов // ФГБОУ ВО «Саратовский ГАУ» им. Н.И. Вавилова. – Саратов, 2017. – 99 с.

Краткий курс лекций по дисциплине «Математическое моделирование технологических процессов» составлен в соответствии с рабочей программой дисциплины и предназначен для обучающихся направления подготовки 35.03.07 Технология производства и переработки сельскохозяйственной продукции. Краткий курс лекций содержит теоретический материал и практические примеры по основным понятиям, определениям и областям применения методов математического анализа и моделирования процессов в сфере АПК. Курс нацелен на формирование практических навыков и умений применять передовые методы анализа и моделирования при решении научно-исследовательских и производственных задач с использованием современных информационных и сетевых технологий.

УДК 004
ББК 32.97

© Розанов А.В., 2017

© ФГБОУ ВО «Саратовский ГАУ», 2017

Введение

Краткий курс лекций по дисциплине «Математическое моделирование технологических процессов» предназначен для обучающихся по направлению подготовки 35.03.07 Технология производства и переработки сельскохозяйственной продукции. Он содержит теоретический материал и практические примеры по основным понятиям, определениям и сферам применения математического моделирования в науке, производстве и переработке сельскохозяйственной продукции. Курс нацелен на формирование ключевых компетенций, необходимых для эффективного решения научно-исследовательских и производственных задач в сфере их профессиональной деятельности с использованием средств компьютерного моделирования и новейших информационных технологий.

Лекция 1

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ СИСТЕМ

1.1. Системный подход. Основные понятия

Управлять современными сложными технологическими процессами уже невозможно на традиционных понятиях «здорового смысла». Интенсивные технологии, опирающиеся на сложные системы машин, средств защиты растений и пр., чрезвычайно расширяют межотраслевые взаимосвязи, обуславливают необходимость одновременного учета взаимодействия большого количества факторов, предвидения отдаленных последствий принимаемых решений. В этих условиях специалист любого уровня должен овладеть системным мышлением.

Новый тип научного мышления связан с **системным восприятием** окружающих нас объектов, явлений, процессов, мира в целом. Независимо от конкретной предметной области исследования, экспериментатор ставит своей конечной целью **управление** теми или иными системами. Чтобы управлять системой, необходимо изучить **свойства** ее поведения. До формирования общей теории систем и кибернетики как науки в исследованиях преобладал метод расчленения сложных явлений на части с целью выявления элементарных законов. Системный подход предполагает изучение сложного явления как целого без дробления на части, исследование, прежде всего, его общесистемных свойств. Исследование систем с позиций управления ими позволило обнаружить сходство и единство процессов управления, происходящих в системах самой различной природы (биологических, технических, экономических).

Суть единства заключается в подобии процессов передачи, хранения, переработки информации. Общими для всех естественнонаучных и социально-экономических наук стали такие фундаментальные понятия, как **система, системные свойства, информация, обратная связь, модель, управление**. Стало возможным опираться на единые математикологические методы. Существенно расширилась область применения формализованных методов научного анализа и синтеза на основе алгоритмизации, программирования, моделирования на ЭВМ.

Системный подход предполагает, прежде всего, рассмотрение любого явления в качестве целостной сложной динамической системы. С этих позиций цель эксперимента — изучение определенных аспектов поведения системы, выявление ее системных свойств. Эффективность экспериментального дела определяется глубиной понимания целостности и целенаправленности поведения системы, учетом сложных взаимосвязей в Иерархической структуре систем, возможностью реализации принципа оптимальности в управлении системой. Поэтому возникает объективная необходимость в органической увязке экспериментальной и производственной деятельности с системным подходом, с более глубоким пониманием самого понятия системы.

1.2. Определение системы

Термин «система» общеупотребителен. Мы говорим: система машин, система севооборотов, система земледелия, система правил дорожного движения, солнечная система, система взглядов и т. д.

Каково научное определение системы?

Система — одно из фундаментальных понятий кибернетики. В первом приближении определение понятия системы связано с наличием некоторого **множества, совокупности элементов**. Но не всякое множество элементов образует систему. Система — не механический набор, а совокупность **взаимодействующих** между собой элементов. Так, взаимосвязи и взаимодействия полей севооборота как системы проявляются в определенном их чере-

довании с учетом последствий предшественников, отдельные машины в системе также взаимосвязаны по своим основным параметрам. В то же время для определения системы недостаточно лишь наличия взаимосвязей и взаимодействия между элементами данного множества. Система обязательно предполагает **целенаправленное взаимодействие**, т. е. некоторую **цель**, как результирующее свойство отношений между элементами исследуемого множества. Взаимодействие элементов системы, направленное на достижение некоторой цели, означает, что она способна реализовать определенную функцию. Для таких сложных систем характерны **информационные процессы управления**. Общие закономерности, описывающие поведение таких систем, принципы организации их структур и внутренних взаимосвязей, особенности их поведения в изменяющейся среде являются предметом изучения теории систем.

Всеобщая взаимосвязь и обусловленность явлений — известное всем объективное свойство материального мира. Очевидно, системы, которые мы исследуем, являются частью этого мира. Как же их вычленять в качестве относительно обособленных систем?

Элементы, относящиеся именно к данной системе, обладают особым характером отношений — **связностью**. Следует подчеркнуть относительность понятия «элемент системы». Данная конкретная система сама может рассматриваться как элемент другой системы более высокого порядка и, с другой стороны, может быть дезагрегирована на составные элементы, как ее подсистемы.

Система всегда рассматривается как некоторая обособленная часть более общего. Элементы системы взаимодействуют между собой несколько иначе, чем с другими системами — система обладает особой связностью, что обуславливает ее обособление. Все, что не относится к данной системе, рассматривается как **внешняя среда** по отношению к системе. Следовательно, с методической точки зрения определение системы означает вычленение ее из среды как некоторой обособившейся части, обладающей особыми свойствами отношений между элементами. В самом характере взаимодействия элементов системы имеет место целесообразность и целенаправленность. Система — не простой набор элементов (объектов, понятий, явлений и т. д.), а упорядоченная совокупность **целенаправленно взаимодействующих элементов**.

Свойство связности проявляется в форме определенной упорядоченности отношений между элементами системы, определенной внутренней структуры. Именно это позволяет выделить систему как некоторую относительно обособленную совокупность. Все, что не относится к данной системе, является **внешней средой** по отношению к ней. Следовательно, определение системы предполагает нахождение соответствующего способа вычленения системы из окружающей среды как некоторой её обособленной части с особыми отношениями между ее элементами.

Итак, в первом приближении систему можно определить как относительно обособленную и упорядоченную совокупность обладающих особой связностью и целенаправленно взаимодействующих элементов, способных реализовать определенные функции.

Однако понятие системы неправомерно ограничивать только материальными объектами. Более глубокое раскрытие содержания системы как научной категории связано с распространением этого понятия и на абстрактные системы: система понятий, система экономико-математических методов, знаков системы и т. п. Более высокий уровень обобщения понятия системы связан с отношением исследователя к исследуемому явлению. В зависимости от цели исследования одно и то же явление может рассматриваться как элемент разных систем. Цель функционирования системы определяется отношением исследователя к решаемой задаче. Например, при изучении вопросов износа деталей, надежности механизмов трактор рассматривается как элемент технической системы, при анализе производительности труда тот же трактор является элементом системы организации производства и т. д. Следовательно, система в этом случае выступает как **метод познания**, как ступень в процессе познания.

Представление об объектах и процессах как о системах — это способ научной абстракции и в зависимости от степени обобщения само определение системы может быть дано на разных уровнях:

- 1) элементарное определение системы как множества взаимосвязанных и взаимодействующих элементов;
- 2) определение системы как методического средства в подходе к решению каких-либо задач;
- 3) определение системы как философской категории познания.

1.3. Основные свойства системы

1. Важнейшим и определяющим свойством системы является ее **целостность**. Свойство целостности обуславливается целесообразным взаимодействием элементов системы в соответствии с общей целью ее функционирования. Система как целое характеризуется множеством свойств, некоторые из которых можно вывести из анализа особенностей отдельных элементов. Но целое всегда обладает и качественно иными свойствами, не выводимыми из свойств отдельных элементов, образующих систему. Проявление качественно новых свойств, не присущих отдельным элементам системы, называют **эмерджентностью**.

Эмерджентность имеет место во всех достаточно больших и сложных системах. Так, лес обладает свойствами, отличными от свойств отдельных деревьев, поведение толпы не является суммой поведений отдельных лиц и т.д. Эмерджентные свойства систем играют огромную роль в управлении ими. Так, сельскохозяйственное предприятие характеризуется системой показателей, отражающих ее своеобразие, свойства. Но исчерпывающая характеристика отрасли в целом не может быть получена только путем статистической сводки системы показателей всех предприятий отрасли. Отрасль как система более высокого порядка обладает новыми, не присущими предприятиям, свойствами, которые должны получить отражение в новой системе показателей. Эмерджентные свойства системы как единого целого обуславливаются проявлением особых эффектов взаимодействия, соответствующей структурой отношений между элементами. Так, эффективность совместного применения удобрений, как правило, выше, чем сумма эффектов от применения отдельных их видов. Эмерджентность является формой реализации некоторых свойств связности и организованности системы.

2. **Связность**. Под свойством связности системы понимается особый характер взаимосвязей между ее элементами. Именно наличие связности является основой вычленения системы из окружающей среды как целого, обособленного. Свойство связности проявляется в форме определенной упорядоченности отношений между элементами, определенной внутренней структуры. При этом понятие структуры должно рассматриваться как вторичное — как отношение отношений между элементами системы. Эффективность функционирования систем существенно зависит от их структуры.

Со свойством связности тесно связано понятие разнообразия системы. Степень разнообразия зависит от числа элементов системы. Максимальное разнообразие определяется как $H = \log_2 N$, где N — число элементов системы. Максимальное разнообразие соответствует тому случаю, когда поведение каждого из элементов не зависит от другого, т. е. элементы не связаны между собой, обладают максимальной степенью свободы. Число независимых характеристик называют числом степеней свободы. Из сущности определения системы следует, что в силу наличия взаимодействия и взаимосвязей между элементами в любой системе число степеней свободы ограничено. Система есть организованная, связанная совокупность элементов и поэтому обладает ограниченным разнообразием. Целенаправленное функционирование системы возможно лишь благодаря ограничению ее разнообразия, что выражается наличием в системе закономерных взаимосвязей. Ограничение степеней свободы, разнообразия лежит в основе управления системой.

3. **Сложность.** Каждой системе свойственна определенная степень сложности, определяемая ее величиной (числом элементов, образующих систему), степенью разветвленности ее внутренней структуры, характером функционирования системы (одноцелевое или многоцелевое функционирование). По числу элементов, внутренней структуре и выполняемым функциям тракторная бригада является менее сложной системой, чем отделение совхоза, в состав которого входит большое число хозяйственных единиц с различными функциями. Научно-технический прогресс в сельском хозяйстве сопровождается нарастанием сложности, которую необходимо преодолевать путем совершенствования структуры и повышением организованности структуры.

4. **Организованность.** Свойство организованности системы проявляется в изменении соотношения между нарастающей сложностью системы и совершенствованием ее структуры. Совершенствование структуры осуществляется путем организации новых форм взаимосвязей и взаимодействий между элементами системы.

Управление системой требует ее соответствующей организации. Благодаря совершенствованию структуры и организованности системы повышается ее управляемость. Таким образом, рассмотренные свойства — **связность, структура, организованность** — являются важнейшими характеристиками, определяющими содержание понятия управления и управляемости систем.

1.4. Система и внешняя среда. Входные и выходные величины

Согласно материалистической диалектике все явления материального мира развиваются в неразрывной связи. Определение системы означает ее распознавание, выделение из окружающей среды как целого, относительно самостоятельного и обособленного, способного реализовать некоторую целевую функцию с рассматриваемой исследователем точки зрения. Вычленение системы из окружающей среды равнозначно разбиению явления на две части: систему и внешнюю по отношению к ней среду. Каждое явление оказывает влияние на другие и само испытывает их ответное воздействие. Поскольку любая система существует в некоторой среде, между средой и системой осуществляются определенные взаимодействия. Так, молочный комплекс как система испытывает влияние таких факторов внешней среды, как состояние кормовой базы, наличие квалифицированной рабочей силы, уровень цен на комбикорма, на основные средства производства и т. д. В свою очередь, комплекс оказывает влияние на среду, потребляя рабочую силу, средства производства, природные ресурсы, поставляя готовую продукцию, изменяя окружающую природную среду и т. д.

Среда оказывает определенные воздействия на систему через соответствующие элементы системы, которые называются входами системы, а факторы внешней среды, осуществляющие эти воздействия, называются **входными величинами**. Так, для вегетирующего зеленого растения входными величинами (факторами внешней среды) являются солнечная радиация, температура окружающего воздуха, наличие углекислого газа и кислорода, почвенной влаги, растворенных в ней элементов минерального питания, различные механические воздействия (ветра, насекомых и т. п.). Эти входные величины оказывают воздействие на систему «растение» через соответствующие элементы системы, образующие вход: хлоропласты листьев (ассимилирующие углекислоту из воздуха и осуществляющие фотосинтез), корневые волоски (всасывающие почвенную влагу с растворенными в ней питательными веществами), покровные ткани всего растения (испытывающие механические, температурные, химические и другие воздействия среды).

Система, в свою очередь, оказывает влияние на среду через определенные элементы, образующие **выход системы**. Факторы, определяющие воздействие системы на среду, называются **выходными величинами**. Выходными величинами системы «растение» являются факторы, определяющие нарастание органической массы, плодоношение, выделение кислорода при фотосинтезе и углекислого газа в процессе дыхания и т. п.

Отдельные производственные процессы также можно рассматривать с точки зрения теории систем. Так, при уборке зерновых на входе уборочного агрегата входными величинами выступают урожай на корню, горючее, смазочные и вспомогательные материалы, потребляемые агрегатом, труд механизаторов; на выходе системы можно наблюдать ряд результативных величин: убранные гектары, намолоченное зерно и солома, экономические показатели работы агрегата.

Факторы, определяемые воздействием входных величин (как внешние воздействия), называют **импульсами**, а выходные величины — **реакциями** системы на соответствующие импульсы. При исследовании системы входные и выходные величины (импульсы и реакции) целесообразно рассматривать как переменные, могущие принимать определенные значения. Таким образом, переменные выступают как количественные характеристики системы.

Переменные, значения которых в пределах данного исследования остаются неизменными, принято называть **параметрами** системы.

Факторы внешней среды представляют собой независимые переменные (обозначаются вектором X), а реакции системы — функции (вектор Y). Выражение входа и выхода системы через переменные позволяет математически формализовать процесс исследования поведения системы, рассматривая ее выходные величины как функции от входных: $Y = F(X)$. Например, урожайность есть некоторая функция от доз внесения удобрений, объем намолоченного за день зерна — функция урожайности, прирост органической массы зеленого растения — функция фотосинтетически активной радиации Солнца и т. п.

Система испытывает влияние бесконечного многообразия факторов внешней среды и соответственно может реагировать на них столь же многообразно. В реальной действительности рассмотрение всех без исключения воздействий и реакций невозможно, т. е. перебор всех возможных переменных системы неосуществим. Обычно задача исследования заключается в рассмотрении того или иного конкретного аспекта явления, характеризующегося ограниченным набором переменных. В этом смысле система представляет собой совокупность наиболее существенных переменных, освещающих данное явление с исследуемой точки зрения. Так, для экономиста совхоза наиболее существенными переменными конкретного работника служат квалификация, стаж работы, образование и т. д., а с точки зрения врача более существенны такие переменные, как возраст, вес, частота пульса, артериальное давление и др. Определение системы как комплекса существенных переменных служит основой математического моделирования ее поведения. Исследование системы предполагает определение ее элементов, выражение их в виде переменных, выбор наиболее существенных из них исходя из цели исследования, нахождение меняющихся значений переменных, выделение параметров.

Глубина исследования системы зависит от степени детализации переменных на входе и выходе. Степень детализации импульсов и реакций называют **разрешающим уровнем исследования** системы. Минимальным является тот разрешающий уровень, когда наблюдатель различает лишь один вход и один выход, хотя предполагается, что в данной системе имеют место разнообразные импульсы и реакции. Например, рассматривая зависимость урожайности от доз удобрений, экономист может ограничиться рассмотрением лишь одной входной величины — общей дозы минеральных удобрений в центнерах действующего вещества в расчете на 1 га посева, тем самым агрегируя все факторы минерального питания растений в одной переменной. Соответственно выходной величиной послужит показатель общего прироста урожая. При более глубоких биологических экспериментальных исследованиях на входе системы в качестве переменных могут быть рассмотрены конкретные формы усвояемого азота, фосфора, калия, микроэлементов и других факторов и, соответственно, на выходе могут различаться реакции растений в виде изменения листовой поверхности, форм ветвления, интенсивности процессов фотосинтеза, роста, развития и т. д.

Вопросы для самоконтроля

1. Что называют системой, элементом системы, структурой системы.
2. В чём состоит системный подход при изучении сложных систем различной природы
3. Основные свойства систем и их взаимосвязь
4. Что принято называть эмерджентностью системы, и каковы её особенности
5. Взаимодействие системы с внешней средой. Основные понятия

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 2

СИСТЕМНЫЙ АНАЛИЗ

2.1. Классификация систем. Виды систем

Элементы, образующие систему, могут быть самыми различными по своей природе. Если система состоит из множества материальных объектов, то данная совокупность элементов представляет собой физическую систему. Так, здание как физическая система состоит из множества конструктивных элементов; физическая система «трактор» представляет собой множество деталей, узлов, механизмов; множество планет и Солнце образуют Солнечную систему и т. п. По специфике составляющих элементов можно различать знаковые системы (система линейных уравнений, азбука Морзе и т. п.), системы понятий (система категорий политической экономии), системы взглядов (философские взгляды Чернышевского), системы правил (правила дорожного движения) и т. д.

Одни из названных систем являются *естественными*, образовавшимися без участия человека, другие создаются людьми и представляют собой *искусственные* системы.

По характеру взаимодействия со средой различают открытые и замкнутые системы. В открытой системе происходит непрерывный обмен с внешней средой энергией, веществом, информацией. Открытая система находится во взаимодействии со средой. В замкнутой системе элементы взаимодействуют только между собой, но не связаны с внешней средой. Строго говоря, таких систем быть не может. Любые системы подвергаются воздействию среды и сами влияют на нее. Но иногда в методических целях полезно абстрагироваться от несущественных с позиций данной задачи взаимодействий системы со средой и рассматривать ее как замкнутую, например, только в информационном отношении. Так, электронная вычислительная машина, выполняющая в автоматическом режиме расчеты по заданной программе, представляет собой информационно замкнутую систему.

Системы различают по характеру причинной обусловленности событий в процессе взаимодействия элементов. Если в процессе взаимодействий последовательность событий в цепи «причина-следствие» однозначно предопределена, то такие системы называют *детерминированными*. Связи в них носят жесткий, функциональный характер. Поведение таких систем в любой момент времени предсказуемо (например, солнечные затмения). С точки зрения процессов управления такие системы не представляют интереса. В то же время следует иметь в виду, что в реальной действительности строго детерминированные системы встречаются редко. Случайность существует всюду, и даже в самых простых детерминированных системах могут произойти случайные сбои. Так, нажатие на спусковой механизм заряженного пистолета сопровождается выстрелом не во всех 100% случаев. Понятие детерминированных систем в этом смысле относительно.

Чаще всего встречаются *вероятностные* системы. Вероятностными называются такие системы, последовательность событий в которых строго не детерминирована, носит вероятностный характер. Поведение таких систем точно предсказать невозможно, оно предсказуемо лишь с определенной вероятностью. Прогнозы о поведении таких систем строят в терминах теории вероятностей. Например: «с вероятностью 0,95 можно утверждать, что прирост урожайности в расчете на 1 ц минеральных удобрений в данном хозяйстве в условиях данного года составит $1,5 \pm 0,4$ ц ». В этой формулировке вероятность 0,95 характеризует степень достоверности вывода (в пяти случаях из ста поведение системы может выйти за указанные пределы 1,1—1,9 ц). Величина $\pm 0,4$ характеризует ошибку прогноза при заданной вероятности.

При рассмотрении свойств систем было отмечено, что системы различаются числом элементов, степенью разветвленности структуры, разнообразием, т. е. сложностью. По степени сложности системы принято делить на:

- 1) простые,
- 2) сложные,
- 3) очень сложные.

Простыми называют системы, состоящие из небольшого числа элементов, с простыми взаимосвязями, неразветвленной внутренней структурой и предназначенные для выполнения элементарных функций. Исследовать и описывать такие системы достаточно легко. Примерами простых систем являются: селекторная связь между двумя абонентами, чередование культур в севообороте и т. п.

Система называется **сложной**, если число элементов в ней значительно, структура взаимосвязей и взаимодействий носит разветвленный характер, выполняемые функции разнообразны. В то же время, несмотря на сложность структуры и функций, система поддается описанию. Например, сельскохозяйственное предприятие можно считать сложной системой.

Очень сложными принято называть системы, сущность взаимосвязей в которых не вполне понятна, недостаточно изучена. Исчерпывающее описание структуры и поведения таких систем при данном уровне знаний не представляется возможным. Очень сложными системами являются общество, экономика, мозг, вселенная.

В кибернетике, где в качестве предмета исследования выступают процессы управления, нет необходимости рассматривать все возможные признаки классификации систем. Основная их классификация здесь строится путем комбинирования двух признаков: сложности и характера причинной обусловленности поведения систем.

Рассматривая приведенную классификацию систем, следует иметь в виду, что пока не существует количественной меры оценки сложности систем по числу элементов, степени связности, характеру структуры, организованности. Поэтому границы приведенной классификации достаточно условны. Жестких критериев дифференциации систем по сложности нет. Условность и относительность классификации состоит еще в том, что иногда термином «сложная система» обозначают не элемент классификации, а признак метода исследования системы при решении многоцелевых задач. Например, при обосновании выбора места для размещения откормочного комплекса приходится рассматривать одновременно несколько несравнимых аспектов: состояние кормовой базы, ресурсы рабочей силы, возможности поставки молодняка, технологию производства, размер капиталовложений, обеспечение водой, охрану окружающей среды и т. д. Проблема сложна в том смысле, что включает разнородные подзадачи, которые должны быть вычленены и изучены. Под **сложной системой** в данном случае понимается признак (или категория) метода разложения проблемы на составные части по специальным аспектам. Решение проблемы состоит в нахождении области пересечения рассматриваемых аспектов.

Другой подобный метод при исследовании систем основан на использовании понятия «большая система». Термин «большая система» не означает, что системы делятся на большие и малые. Под **большой** системой понимается методическое средство, используемое при анализе систем, которые невозможно охватить в целом, поскольку они необозримы либо в пространстве, либо во времени и поэтому исследуются по частям (не по аспектам, как в сложной системе).

Таким образом, понятия «большая система» и «сложная система» с методической точки зрения представляют собой средства, позволяющие осуществлять разные способы разложения системы на подсистемы: в «большой системе» на подсистемы расчленяется объект исследования, а в «сложной системе» на подсистемы дифференцируется задача исследования.

2.2. Структура систем. Иерархические структуры

При рассмотрении основных свойств систем отмечалось, что важнейшей особенностью всех систем является наличие некоторой структуры как упорядоченной формы отношений взаимодействия. Для сложных систем, рассматриваемых в кибернетике, характерны **иерархические** (многоуровневые) структуры. Принимая любой элемент системы в качестве от-

правного уровня, его можно рассматривать как подсистему вышестоящего уровня и, наоборот, элементы данного уровня можно рассматривать как систему, включающую подсистемы нижестоящего уровня. Уровни иерархии могут различаться по различным признакам: по организационному признаку (определение субординации при решении управленческих задач), по тому или иному аспекту деятельности системы (уровни связи по технологии производства, по со подчиненности отраслей), по способу расчленения сложной проблемы на иерархию более простых задач, по временным интервалам и т. д. В каждом из названных случаев образуются многоуровневые *вертикальные или горизонтальные иерархические структуры*.

Иерархические структуры широко распространены и носят универсальный характер, что объясняется их существенными преимуществами по сравнению с другими типами структур (например, радиальными). Наличие многоуровневой иерархической структуры обеспечивает системе высокую надежность функционирования благодаря возможности создания элементной избыточности. Экономичность системы обеспечивается благодаря рациональной дифференциации энергетических, материальных и информационных потоков по уровням. При иерархической структуре создается возможность целесообразного сочетания локальных критериев с общим критерием при нахождении оптимального режима функционирования системы. Иерархические структуры широко используются в системах управления производством. Эффективность функционирования системы в значительной мере зависит от ее структуры, от формы связей в иерархии. Поэтому одной из важных задач кибернетики является изучение закономерностей в иерархических структурах.

2.3. Состояние и движение системы

При рассмотрении отношений системы со средой отмечалось, что входные и выходные величины выступают как переменные, могущие принимать те или иные значения. Найти конкретные значения наиболее существенных переменных системы — значит охарактеризовать состояние системы.

С течением времени значения переменных изменяются (элементы системы из одного состояния переходят в другое). В некоторых системах переход из одного состояния в другое осуществляется мгновенными скачками, а затем система относительно долго остается в данном состоянии. Такие системы называют *статическими*. В реальной действительности система не может оставаться в статическом состоянии бесконечно долго. Все системы изменяются во времени. Когда говорят о статических системах, имеется в виду, что существенные переменные системы остаются неизменными в течение времени данного исследования.

Наибольший интерес с точки зрения управления представляют закономерности поведения сложных динамических систем. Системы называются *динамическими*, если переход из одного состояния в другое совершается не мгновенными скачками, а в течение некоторого времени, т. е. процесс перехода можно наблюдать и описать.

Последовательное изменение состояний системы называется *движением*. Движение системы в широком смысле слова включает в себя любое ее изменение во времени. Движение системы описывается значениями ее переменных в последовательные моменты времени. Если состояние системы в момент времени t_0 описывается вектором X_0 , а состояние в момент t_1 вектором X_1 то, следовательно, система совершила переход из состояния

$$t_0 (X_0) \text{ в } t_1 (X_1).$$

Переход системы из одного состояния в другое обычно совершается по некоторому закону, строгое описание действия которого называют *оператором*. Переменная, подвергающаяся влиянию оператора, называется *операндом*, а новое значение переменной — *образом*. Так, под воздействием труда и средств производства исходное сырье или природный ма-

териал превращаются в потребительскую продукцию (здесь операнд — сырье, образ — продукция, оператор — технологический процесс, совершающийся с определенной закономерностью).

Для описания состояния системы и ее движения используются различные способы: словесный, табличный, матричный, графический. Иногда удается найти обобщенное математическое выражение, описывающее закономерность движения системы в аналитической форме. При описании состояния и движения хозяйственных систем чаще всего используются графический, табличный и матричный методы.

Переменные, характеризующие состояние системы, можно представить в виде координат в пространстве. Если исследуются только две переменные системы, то ее состояние будет охарактеризовано точкой в двумерном пространстве (на плоскости). Обычно число переменных значительно больше и поэтому состояние системы характеризуется точкой в многомерном векторном пространстве. Такое пространство, в котором каждому состоянию системы соответствует определенная и причем единственная точка, называется **пространством состояний системы**.

В общем случае число возможных состояний системы определяется числом элементов, их разнообразием и связями между ними. Допустим, система представлена отделением совхоза и характеризуется тремя переменными: y_1 — урожайность зерновых культур, ц/га; y_2 — удой на корову за год, кг; y_3 — количество работников.

Состояние системы будет изображаться точкой в трехмерном пространстве. В реальных условиях каждая переменная изменяется в ограниченном интервале.

Область пространства, в которой может находиться точка, изображающая состояние системы, называется **областью допустимых состояний**. Эта область включает в себя всевозможные траектории поведения системы.

В процессе движения системы один и тот же оператор обычно оказывает воздействие на множество операндов. Например, известкование почвы оказывает влияние не только на кислотность почвенного раствора, но и на почвенную микрофлору, механический состав почвы, поглотительную способность и т.д.; дополнительные капитальные вложения воздействуют на обеспеченность основными фондами, вооруженность труда, занятость рабочей силы, сезонность производства, производительность труда, рентабельность производства и другие переменные предприятия, являющиеся в используемой здесь терминологии операндами. Новые значения перечисленных переменных, получаемых под влиянием возрастающих капиталовложений, образуют множество образов. В качестве оператора следует рассматривать сложную и не всегда достаточно изученную закономерность изменения множества переменных под влиянием дополнительных капитальных вложений.

Переход множества операндов во множество образов под воздействием некоторого оператора называется **преобразованием**. Преобразования могут быть однозначными, взаимно однозначными, однозначными односторонне и неоднозначными. Если в преобразовании каждому оператору соответствует только один образ, то такое преобразование называется однозначным.

Исследование поведения систем в динамике позволяет раскрыть суть оператора преобразования — найти закономерность движения системы.

Некоторые преобразования формально действительны и в обратном направлении: зная закон преобразования, например, в рядах динамики, можно вычислить не только последующие, но и предыдущие уровни урожайности. Следовательно, зная оператор, по образу можно восстановить значение операнда. Такие преобразования называются взаимно однозначными. Некоторые преобразования могут быть лишь односторонне однозначными, когда один и тот же образ может получиться как результат перехода различных операндов. Например, повышение урожайности с 1 га до 45 ц может быть результатом повышенных доз удобрений, полива, применения нового сорта и т. д.

Движение систем обычно описывается табличным, графическим, матричным или другими способами. Для изображения движения технических систем чаще используются кинематические графики.

В реальных условиях преобразования, как правило, неоднозначны. Так, дополнительное внесение одной и той же дозы удобрений в различные годы и на разных типах почв дает разную прибавку, а иногда может сопровождаться и снижением урожайности (например, при засухе). Даже при хорошо выровненных условиях эксперимента прирост урожайности бывает различным. Повышение уровня механизации может привести к снижению себестоимости продукции, но не всегда. Механизация может просто облегчить труд работников. Описанные случаи являются примерами неоднозначных преобразований.

При исследовании взаимно неоднозначных преобразований важно не только определить возможное разнообразие появляющихся образов, но и установить вероятности перехода данного операнда в тот или иной образ (например, вероятность получения определенного числа всходов от 100 семян, вероятность изготовления детали с браком и т. д.). В сельском хозяйстве после благоприятного по погодным условиям года может снова следовать благоприятный год, но не исключено, что он окажется неблагоприятным или средним. Вероятности таких переходов можно оценить на основе статистической обработки рядов динамики урожайности сельскохозяйственных культур.

2.4. Методы изучения состояния и движения системы

Движение системы, как видно из изложенного выше, есть процесс преобразования входных величин в выходные. Иначе говоря, импульсы на входе системы ($X_1, X_2, \dots, X_n$) соответственно преобразуются в реакции на выходе системы ($Y_1, Y_2, \dots, Y_m$). При этом реализуется некоторая закономерность $Y = F(X)$. Сама система в данном случае представляет собой некоторый преобразователь входных величин в выходные. С этой точки зрения производство есть преобразователь ресурсов (природных ресурсов, основных и оборотных средств, ресурсов труда и т. п.) в производимую продукцию.

Рассматривая любую систему как некоторый преобразователь, ее поведение можно изучить путем анализа входных и выходных величин. При этом о поведении системы делается вывод только на основе исследования импульсов и реакций, а внутренняя структура системы и протекающие в ней процессы не рассматриваются. Такой способ изучения поведения системы получил название **метода черного ящика**. Метод используется чаще всего в тех случаях, когда внутренняя структура исследуемой системы либо очень сложна и имеет многоуровневую иерархию, а эмерджентные свойства системы в целом невозможно вывести из свойств отдельных подсистем, либо внутренние свойства системы вовсе не известны. Тогда исследователю остаются доступными лишь вход и выход системы. Оказывая на вход системы разнообразные воздействия, исследователь изучает ее реакции на выходе. Таким методом система может быть изучена в достаточной степени, что делает возможным предсказание ее поведения при любых импульсах на входе, хотя внутренняя структура системы при этом специально не исследуется.

В то же время следует иметь в виду, что одинаковым поведением могут обладать системы, имеющие совершенно разную структуру. Это открывает возможность изучения поведения менее доступных систем на более доступных и простых по структуре. На этом принципе основан другой широко используемый метод кибернетики — **моделирование**.

Процесс моделирования заключается в том, что реальное явление и его аналог (физическая, математическая или иная по своей природе модель явления) рассматриваются как два черных ящика, между входами и выходами которых существует однозначное соответствие. Это значит, что вход второго черного

ящика может быть однозначно преобразован во вход первого и соответственно выход второго может однозначно преобразовываться в выход первого. Это явление получило название **изоморфизма**. Изоморфными относительно друг друга являются записи чисел на

перфокартах, перфолентах, магнитных лентах, поскольку они тождественны по содержанию. Метод моделирования по принципу черного ящика основан на явлении изоморфизма, т. е. на взаимно однозначном соответствии поведения некоторых объектов независимо от их внутренней структуры.

Следует, однако, иметь в виду, что реальное явление всегда богаче, разностороннее, чем модель, на которой исследуются лишь наиболее существенные переменные реальной системы. Так, на географической карте отображаются не все детали местности, а лишь наиболее важные; в экономико-математической модели оптимизации производства учитываются только основные условия хозяйства и т. д. По модели невозможно во всех деталях создать заново моделируемый объект, хотя в самом существенном их поведения тождественны. Каждому состоянию реальной системы однозначно соответствует определенное состояние модели. Но в модели отображаются лишь существенные переменные, а в реальной системе переменных значительно больше. Поэтому состояние реальной системы не может определяться данным состоянием модели. Следовательно, в моделировании мы имеем случай соответствия, однозначного лишь в одну сторону.

Отношения подобия, однозначного лишь в одну сторону, называются *гомоморфизмом*. В предельном случае гомоморфизм становится изоморфизмом. В кибернетике используются гомоморфные модели. Так, экономико-математическая модель является гомоморфной моделью какого-либо экономического процесса.

Для сложных кибернетических систем связь между входными и выходными величинами, как правило, опосредована состоянием внутренних параметров. Поэтому при изучении поведения системы в целях совершенствования управления приходится рассматривать состояние внутренних ее параметров. Рассмотрим, например, работу уборочного агрегата. Выходными величинами Y в данном случае будут количество убранных гектаров y_1 , масса намолоченного зерна y_2 , высота среза стеблей y_3 , количество зерен, остающихся невымолоченными в колосьях, y_4 , количество неподобранных стеблей y_5 и т. д.

Входными величинами являются воздействия комбайнера на те или иные механизмы комбайна (изменение положения штурвала x_1 , рычага коробки передачи x_2 , регулятора подачи воздуха x_3 , изменение расстояния между барабаном и декой x_4 и т. д.). На агрегат оказывает влияние также ряд внешних воздействий среды (неровности рельефа x_5 , густота и высота стеблей x_6 , направление и сила ветра x_7 , влажность растений x_8 и др.).

Входные величины X вначале преобразуются в координаты самой системы (агрегата). Обозначим эти координаты через Z . Координатами системы являются, например, число оборотов двигателя z_1 , высота режущего аппарата z_2 , расстояние между мотовилом и режущим аппаратом z_3 , скорость z_4 , ширина зазора между барабаном и декой z_5 и т. д.

Координаты системы Z , в свою очередь, преобразуются в выходные величины Y как конечные результаты процесса. Преобразование входных величин X в координаты системы Z можно выразить соотношением $Z = F(X)$. Так, число оборотов двигателя z_1 определяется положением заслонки подачи воздуха x_3 , т. е. $z = f(x_3)$; высота режущего аппарата z_2 зависит от положения штурвала x_1 т. е. $z_2 = f(x_1)$ и т. д.

Преобразование координат системы Z в выходные величины Y выражается как $Y = F(Z)$. Так, вымолот зерна из колосьев y_4 определяется такими координатами системы, как z_1 и z_5 , т. е. $y_4 = f(z_1, z_5)$; количество убранных гектаров y_1 зависит от скорости z_4 , т. е. $y_1 = f(z_4)$, высота среза (высота стерни) y_3 определяется высотой режущего аппарата z_2 , т. е. $y_3 = f(z_2)$ и т. д.

Следует также иметь в виду, что выходные величины зависят не только от контролируемых воздействий на систему, но и от случайных воздействий среды (так называемые возмущающие воздействия). Высота среза, например, зависит не только от положения штурвала, определяющего высоту режущего аппарата от горизонтальной поверхности, но и от неровностей рельефа, вибрации аппарата и т. п. Возмущающие воздействия бывают не только внешними. Они могут возникать и в самой системе (например, затупление ножей режущего аппарата и т. п.).

Таким образом, если весь процесс, описанный здесь, рассматривать как черный ящик, то он как бы расчленяется на две части, соотношения между входами и выходами которых последовательно связаны: (1) $Z=F(X)$ и (2) $Y=F(Z)$, т.е. выход первой части служит входом для второй части.

2.5. Типы поведения динамических систем. Понятие устойчивости

В процессе движения динамическая система может находиться в трех различных состояниях: равновесном, переходном и периодическом.

Состояние динамической системы называется равновесным, если не одна из существенных ее переменных не изменяется во времени. Различают статическое и динамическое равновесие. Статическое равновесие — это равновесие покоя. Более важным с точки зрения процессов управления является равновесие систем в динамике.

Равновесное состояние системы в динамике означает, что преобразования ее переменных являются тождественными, т. е. не порождают новых образов. Так, на одном и том же уровне поддерживаются страховые запасы семян, кормов, горючего. **Равновесный режим** поведения динамической системы тесно связан с понятием устойчивости системы, о чем будет сказано позже.

Переходным называется такой тип поведения системы, когда она находится в состоянии перехода из некоторого начального состояния в какое-либо установившееся состояние. Переходный режим является основным типом поведения для динамических развивающихся систем, особенно часто встречающихся в экономике.

Если система через равные интервалы времени возвращается в одни и те же состояния, то данный режим поведения называется **периодическим**. Так, известно, что на величину отклонений урожайности от линии регрессии, отражающей основную тенденцию роста, существенное влияние оказывают повторяющиеся с 11-летней периодичностью изменения солнечной активности.

Понятие устойчивости. Важнейшей характеристикой поведения динамических систем является их устойчивость. Под **устойчивостью** понимается постоянство состояния системы или постоянство последовательности некоторых ее состояний во времени в процессе ее преобразований. Понятие устойчивости в живых организмах связывают с понятием гомеостазиса. **Гомеостазисом** называется способность живых организмов обеспечивать оптимальный режим внутренней среды путем поддержания постоянства существенных переменных (температуры, состава клеточной жидкости и т. д.). Обеспечение постоянства существенных параметров системы достигается за счет процессов управления (саморегулирования) в живых организмах, благодаря чему устраняются последствия возмущающих случайных влияний среды на отдельные подсистемы организма.

Принцип гомеостазиса в настоящее время часто используется при конструировании различных технических систем, которые называют гомеостатическими. Первое такое техническое устройство было создано английским нейрофизиологом У. Эшби, которое было названо гомеостатом Эшби. Использование принципа гомеостазиса позволило создать разнообразные технические системы (например, для определения оптимальных значений параметров автопилотов в современных самолетах, систем автоматического регулирования методом случайного поиска и т. п.). Понятие устойчивости в строгом смысле относится не к самой системе, а к какому-нибудь свойству ее поведения, причем поведения системы в целом, а не какой-либо ее части. При агрегировании ряда систем с неустойчивым поведением новая система может оказаться устойчивой. На устойчивость системы может оказать влияние также способ агрегирования. В то же время при объединении нескольких устойчивых систем не обязательно получится система с устойчивым поведением.

Понятие устойчивости динамических систем обычно связывается с тем, что отклонения некоторых параметров системы от заданных не превышают допустимых значений. Например, отклонения урожайности от линии тренда, как правило, находятся в определенных гра-

ницах, что свидетельствует об устойчивости основной тенденции ряда, обусловливаемой факторами интенсификации. Устойчивость основной тенденции является важным условием, обеспечивающим надежное прогнозирование поведения системы.

Следует заметить, что система, устойчивая по некоторому свойству поведения, по другому признаку может быть и неустойчивой. Так, устойчивость основной тенденции в ряду динамики урожайности вовсе не означает, что окажется устойчивой также тенденция изменения себестоимости продукции. Определенный интерес представляет исследование устойчивости поведения систем по отношению к возмущениям различной природы (например, устойчивость тенденции роста урожайности в условиях орошения, при различных сочетаниях погодных условий по зонам страны и т. п.).

2.6. Кибернетические системы

Кибернетика как наука рассматривает не любые системы, а только определенный класс систем, называемых **управляемыми системами**.

Основной особенностью управляемых систем является их способность изменять свое поведение под влиянием управляющих воздействий на вход системы. Свойство управляемости присуще объектам самой различной природы (технические, биологические, экономические и т. д.). Конкретные особенности этих систем, зависящие от их природы, не являются предметом рассмотрения кибернетики. Здесь представляют интерес общие закономерности изменения поведения управляемых систем под влиянием различных управляющих воздействий. С этой точки зрения объектом исследования является абстрактная кибернетическая система. При рассмотрении кибернетической системы мы не принимаем во внимание материальные и энергетические взаимосвязи системы со средой. Нас здесь, прежде всего, интересуют информационные потоки, а материальные и энергетические потоки выступают лишь как носители информации. Такой подход позволяет вычленивать и исследовать специфические процессы управления, протекающие в системе. В связи с этим такая абстрактная кибернетическая система получила название **системы управления**.

Система управления имеет особую структуру. В простейшем случае она состоит из двух подсистем: а) управляющей и б) управляемой. Эти подсистемы соединены каналами связи. С точки зрения управления системой, важно рассмотреть, прежде всего, информационные процессы. Управляющая подсистема в зависимости от цели управления оказывает определенные воздействия на вход управляемого объекта. Линия, по которой осуществляется это воздействие, называется **каналом прямой связи**, а сама форма связи — **прямой**.

Под влиянием управляющих воздействий исходное состояние управляемой подсистемы преобразовывается в новое состояние. Цель управляющих воздействий — привести систему в желаемое состояние, а если речь идет об управлении поведением системы, — то добиться соответствия движения управляемого объекта по некоторой траектории в соответствии с его функциями. В этом смысле управление есть процесс, ведущий к достижению некоторой цели, т. е. заставляющий систему выполнять определенную программу.

Управляемый объект всегда испытывает влияние ряда неконтролируемых управляющей подсистемой случайных воздействий среды. В связи с этим выходные величины могут отклоняться от требуемых значений. Если управляющая подсистема не будет знать об этих отклонениях, то фактическое поведение системы очень скоро отклонится от заданной траектории настолько, что достижение цели управления станет невозможным. Поэтому очень важно, чтобы информация о выходных величинах (или о величине их отклонений от заданных) систематически поступала в управляющую подсистему, т. е. необходим канал связи, соединяющий выход системы с входом. Эта линия называется **линией обратной связи**, а информационное воздействие выхода на вход системы называется **обратной связью**. Благодаря наличию обратной связи создается возможность постоянно сопоставлять выходные величины с требуемыми их значениями и в зависимости от величины отклонений принимать новые управленческие решения. При этом, естественно, управляющая подсистема должна не толь-

ко учитывать отклонения выходных величин от заданных, но и принимать во внимание изменения условий внешней среды и оценивать возможное влияние этих изменений на поведение управляемого объекта.

Обратная связь в системах управления носит универсальный характер. Она широко распространена в технике, природе, обществе. Отдельные элементы подсистем могут иметь местные обратные связи. Обратная связь может осуществляться либо непосредственно между выходом и входом, либо через опосредствующие элементы системы. Между выходом и входом системы как единого целого образуется контур **внешней обратной связи**.

Если целью управляющих воздействий является устранение последствий влияния возмущающих факторов внешней среды на систему, то управляющие воздействия обычно носят компенсирующий характер. Суммирование этих воздействий с величинами отклонений от нормы должно дать нулевое значение, т. е. ликвидировать отклонение. Такая обратная связь называется **отрицательной**. **Отрицательная обратная связь стабилизирует поведение системы**. Для систем с отрицательной обратной связью характерно состояние равновесия (статического или динамического). Обратная связь называется **положительной**, если она способствует увеличению отклонения выходной величины от начального уровня. Например, материальное стимулирование способствует дальнейшему росту производительности труда, увеличению объема продукции.

Управляющая и управляемая подсистемы имеют существенные различия. **Управляющая подсистема** есть не что иное, как преобразователь информации о состоянии управляемой подсистемы и внешней среды в управляющие воздействия. В экономических системах эти управляющие воздействия носят характер директивных указаний, распоряжений, предписаний, которые призваны ограничивать множество возможных состояний управляемой подсистемы, т. е. осуществлять целенаправленный выбор некоторого типа поведения системы в целом.

В управляемом объекте происходит преобразование ресурсов в продукты потребления (если речь идет о производственных объектах). Информация о ходе этих процессов снимается на выходе и снова поступает в управляющую подсистему. Управляющая подсистема снова вырабатывает сигналы воздействия, которые корректируют поведение управляемого объекта. Большое значение имеет качество управления. Величина, позволяющая оценить качество управления, называется **критерием эффективности управления**.

Например, если целью управления является стабилизация системы, то чем меньше отклоняются выходные величины от заданных, тем выше качество управления. Если управляющие воздействия обеспечивают наибольшее значение критерия эффективности, то управление является **оптимальным**.

Отличительная особенность производственных систем управления заключается в том, что это так называемые **человеко-машинные системы**. Человек определяет цель функционирования системы, является одновременно и носителем, и пользователем информации, выработка управленческих решений также осуществляется человеком. Человеко-машинные системы управления имеют специфические черты, изучение которых требует привлечения методов ряда смежных дисциплин: биологической кибернетики, технической кибернетики, экономической кибернетики. Важной задачей исследования человеко-машинных систем является выяснение вопросов оптимального распределения функций между человеком и машиной с учетом их возможностей и особенностей.

2.7. Особенности систем в сфере пищевых производств

Отличительная особенность пищевых производств в сфере АПК — переплетение технологических процессов с биологическими. Знание закономерностей развития биологических систем имеет существенное значение в повышении эффективности сельскохозяйственного производства. Современные методы исследования, разрабатываемые в кибернетике, способствуют углублению знаний о процессах саморегулирования в биологических системах.

Хотя процессы управления в живой природе, обществе и технических устройствах бесконечно разнообразны, они имеют общие черты, общие закономерности. Одним из основных достижений кибернетики как науки является выяснение того факта, что закономерности процессов управления идентичны как для естественных, так и для искусственных систем. Основоположник кибернетики Норберт Винер определил ее как науку «об управлении и связи в животном и машине». Классические методы исследования в биологии значительно углубились благодаря новому подходу на основе системного анализа. В настоящее время развивается биологическая кибернетика как самостоятельная отрасль кибернетики.

Биологическая кибернетика изучает общие принципы и механизмы целесообразного саморегулирования в живых организмах и биологических сообществах в условиях изменяющейся окружающей среды. В отличие от классических методов исследования в биологической кибернетике живые организмы рассматриваются с точки зрения общей теории систем, изучения их системных свойств. На первое место выдвигаются вопросы исследования процессов управления (саморегулирования) в живых организмах и информационные процессы.

Важнейшим методом изучения биологических систем становится **моделирование**. При этом используются математические методы и электронно-вычислительная техника.

Биологические системы характеризуются сложной многоуровневой иерархической структурой. Например, если в качестве отправного уровня рассматривать биоценоз (совокупность систем организмов), то иерархия биологических систем может быть представлена следующими уровнями:

Биоценоз → вид → организм → орган → ткань → клетка → субклеточные структуры

Биологические системы имеют существенные особенности. Фундаментальной особенностью биологических систем является их **способность к самоорганизации**, что включает в себя избирательное поглощение вещества и энергии из окружающей среды и осуществление непрерывного активного обмена веществ. Только живым системам присущи такие свойства, как раздражимость, размножение, обмен веществ.

Основная форма управления в биологических системах — **целесообразная саморегуляция**. Оптимизация поведения биологических систем в изменяющейся окружающей среде осуществляется в форме саморегулирования. Целесообразность поведения биологических систем заключается в их стремлении к повышению организованности, совершенствованию внутренней структуры, надежности функционирования системы. Встречаются различные типы саморегулирования биологических систем.

При появлении возмущающих воздействий факторов внешней среды биологическая система способна поддерживать некоторые свои параметры на постоянном уровне. Таким образом, биологическая система обладает свойством **стабилизации**.

Биологические системы можно рассматривать как ограниченно открытые. Входы и выходы биологических систем контролируются самой системой в зависимости от ее состояния (например, уменьшение испарения за счет регулирования механизма устьиц).

Структура саморегулирования биологических систем специфична. Управляющая и управляемая подсистемы в них не всегда четко разграничены, они переплетаются между собой. При длительных воздействиях возмущающих факторов внешней среды биологическая система способна перестраивать свою внутреннюю структуру и изменять некоторые функции, т. е. она обладает свойством **адаптации** к изменяющимся условиям внешней среды.

Существенной особенностью биологических систем является их способность к **эволюции** при систематических изменениях окружающей среды. Благодаря развитию новых форм биологические системы приспособляются к изменяющимся условиям в такой степени, что становятся способными использовать новые виды вещества и энергии из окружающей среды.

В отличие от технических систем биологические обладают способностью к **самовосстановлению** (созданию новых структур взамен разрушенных жесткими факторами внешней среды) и **самовоспроизводству** (размножению).

Механизм обратной связи в биологических системах также специфичен. В процессе эволюции, а также под влиянием мутаций происходят определенные изменения наследственности. В процессе размножения эти новые свойства реализуются в новых формах. Если последние оказываются более приспособленными, чем существующие ранее, то они размножаются более быстрыми темпами. Следовательно, с кибернетической точки зрения реализуется положительная обратная связь. Если же изменившиеся особи оказываются менее приспособленными, то они гибнут, следовательно, срабатывает механизм отрицательной обратной связи, а система стабилизируется на исходном состоянии.

Биологические системы способны к стабилизации, но они, по-существу, являются динамическими, и понятие устойчивости применительно к ним имеет свои особенности. Стабильность биологических систем не означает постоянного равновесия их состояния. Наиболее характерным типом поведения биологических систем является переходный режим. При характеристике поведения биологических систем принято использовать термин «устойчивое неравновесие». Живые организмы никогда не бывают в равновесии в строгом смысле. Благодаря наличию свободной энергии, свойствам раздражимости, возбудимости, сопровождающимся энергетическими процессами, систематическому приспособлению к изменяющимся условиям окружающей среды, равновесие биологической системы постоянно нарушается.

Вопросы для самоконтроля

1. Принципы классификации естественных и искусственных систем
2. Дифференциация больших и сложных систем
3. Статические и динамические системы
4. Методы исследования состояния и движения систем
5. Понятие устойчивости системы

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>

2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>

3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд. – Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>

4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд. – Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>

2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16.

(обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>

3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд., стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>

4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд. – Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 3

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ НА ОСНОВЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

В настоящее время трудно представить себе область человеческой деятельности, где нельзя было бы использовать ЭВМ. ЭВМ применяются буквально во всех сферах нашей жизни: в науке, производстве, культуре, медицине, искусстве и т.д. Уже сейчас персональные компьютеры становятся такими же постоянными спутниками нашей жизни, как наручные часы! Особенно велика роль ЭВМ в инженерной практике, когда инженеру при разработке новых изделий или технологий приходится выполнять большой объем вычислений, оценок, расчетов и других операций по обработке данных.

Применение ЭВМ, и особенно персональных (ПЭВМ), должно в недалеком будущем заметно сократить сроки разработок, снизить стоимость проектирования, экономических расчетов и научно-технических исследований.

3.1. Этапы решения практических задач на ЭВМ

1 этап: Постановка задачи в реальных объектах. Формулировка или осознание проблемы. Составление технического задания.

В инженерной практике это, например, необходимость определения выдержит ли мост данную нагрузку или трубопровод требуемое давление пара.

2 этап: Построение математической модели исследуемого объекта или процесса. На этом этапе выделяют наиболее важные черты и свойства объекта (неважные отбрасывают) и эти свойства записывают в виде математических уравнений.

3 этап: Выбор алгоритма решения и разработка программы для ЭВМ.

4 этап: Выполнение программы на ЭВМ с входными данными, отвечающими реальному процессу.

5 этап: Проверка соответствия полученных результатов требованиям задания или проверка точности решения.

Обычно на практике с первой попытки получить удовлетворительное решение не удается. Поэтому этапы построения модели, разработки алгоритма и программы, выполнения программы и оценки точности результатов повторяют до тех пор, пока не появится уверенность, что найдено наилучшее решение.

6 этап: Составление технической документации, описывающей проектируемое изделие или технологический процесс.

3.2. Основные понятия моделирования

Модель (от латинского *modulus* - мера, образец, способ) - это устройство, воспроизводящее (имитирующее) строение или действие другого ("моделируемого") устройства в научных, производственных (при испытаниях) или спортивных (авиамоделизм) целях.

В широком смысле **модель** - это любой мысленный или условный образ (изображение, описание, схема, чертеж, график, план, картина и т.д.) некоторого объекта, процесса или явления ("оригинала" данной модели), который используется в качестве "заместителя" этого объекта.

Оригинал (от латинского *originalus* - первоначальный) - подлинник, подлинное произведение (в отличие от копии).

Моделирование - исследование явлений или процессов путем построения и изучения их моделей. Моделирование состоит в выявлении основных свойств исследуемого процесса,

построении его модели, изучении свойств (поведения) модели и прогнозировании поведения исследуемого процесса на основе изучения модели.

Критерием правильности моделирования является практика.

3.3. Общая классификация моделей

Вообще говоря, моделей очень много: знаковые, лингвистические, социологические и т.д. Но для инженерной практики наиболее существенны физические и математические модели.

Физической называют модель, воспроизводящую изучаемый оригинал с полным или частичным сохранением его физической природы и геометрического подобия.

Физические модели бывают натурные (1:1) и масштабные (уменьшенные или увеличенные).

Создание физических моделей - дело очень дорогостоящее, особенно для моделирования больших систем и сложных процессов. Однако физическое моделирование - это единственный способ проникновения в неизвестное в тех случаях, когда человеческие знания пока не могут быть теоретически обобщены. В инженерной практике все более эффективным и дешевым становится математическое моделирование.

Математическая модель - приближенное описание явлений внешнего мира, выраженное с помощью математической символики. (Академик А.Н.Тихонов. Математическая энциклопедия).

Математическое моделирование - это мощный инструмент познания мира, а также прогнозирования и управления. Но чтобы описать какое-либо явление природы, надо выявить его наиболее существенные свойства, закономерности и связи, отбросив менее существенные. Это упрощает и огрубляет модель, но позволяет сделать ее доступной для исследования, т.к. в реальном явлении участвует множество процессов, учесть которые просто невозможно, а в конкретной задаче и не нужно. Например, при изучении колебаний обычного маятника (груза на нити) совершенно не важно белая нить или черная, а груз красный или синий - важны длина нити и масса груза.

После того, как наиболее важные факторы выявлены, следующий шаг состоит в описании этих факторов и их взаимосвязи на языке математики (при помощи математической символики) обычно в форме уравнений и неравенств. Это не случайно, т.к. все законы природы сформулированы на языке уравнений: алгебраических, дифференциальных, разностных и т.п.

Когда явление природы, науки или техники переведено на язык математики, оно может быть исследовано хорошо разработанными математическими методами.

Эксперимент на математической модели, как правило, проще, быстрее, дешевле, экономичнее и безопаснее, чем на физической модели, а тем более на оригинале. Более того, хорошая математическая модель может быть использована для изучения многих других оригиналов с различной физической природой, поскольку "математика - это искусство называть разные вещи одними именами" (Анри Пуанкаре).

3.4. Классификация математических моделей

Все разнообразие математических моделей может быть разделено на два больших класса:

1. Детерминированные модели.

2. Вероятностные (эмпирические, статистические, иногда стохастические) модели.

Детерминированными называют математические модели, в которых все определено (от латинского слова *determino* "определяю"), т.е. установлено однозначное функциональное соответствие между входными и выходными параметрами. Главное здесь состоит в том, что для моделирования процессов в таких системах используются обычные уравнения: алгебраические, дифференциальные, в частных производных и т.д. Уравнения могут быть очень

сложными, нелинейными, трансцендентными, но они не должны содержать никаких случайных переменных.

Вероятностными называют модели, в которых соответствие входных и выходных переменных носит случайный характер и описывается распределением вероятностей. Для их исследования приходится пользоваться методами теории вероятностей и математической статистики. Это прежде всего обработка экспериментальных данных. К вероятностным моделям примыкают так называемые имитационные математические модели (имитация - подражание, подделка).

Имитационное моделирование - это воспроизведение в модели процессов, происходящих в оригинале, путем искусственной имитации случайных величин, от которых зависят эти процессы. Случайные величины в модели генерируются с помощью датчиков или генераторов (псевдо)случайных чисел, реализованных в виде специальных программ на ЭВМ. Выходные характеристики модели оцениваются методом **Монте-Карло**. Эти модели применяются в тех случаях, когда нельзя составить и решить алгебраически или численно уравнения, описывающие функционирование исследуемой системы. Яркий пример имитационного моделирования действительности - это компьютерные игры!

3.5. Оптимизационные математические модели

Рассмотренные выше математические модели, имея множество отличий, обладают одной общей чертой. Они предназначены для математического описания различных природных явлений, технологических процессов и социальных феноменов. Это так называемые описательные или, говоря научно, дескриптивные математические модели (от англ. *Description-описание, изображение*). Они служат для познания окружающего мира и овладения его законами.

Не менее, а часто более, важную роль и в жизни, и в инженерной практике играет еще один класс математических моделей, предназначенных не только для описания, но и для управления природными и технологическими процессами в интересах человека и общества. Эти модели называются оптимизационными (от латинского *optimum-наилучшее*).

Оптимизация вообще - это поиск некоторого оптимума - совокупности наиболее благоприятных условий (от латинского *optimum* - наилучшее). В математическом моделировании - это поиск наилучшего варианта решения задачи с точки зрения достижения намеченной цели. Например, создание нового изделия с более высоким качеством и меньшей себестоимостью.

Оптимизационными называются такие математические модели, которые позволяют найти наилучшее решение поставленной задачи или наилучший способ достижения цели при заданных условиях и ресурсах.

(Ресурсы-запасы, средства, возможности, источники средств, доходов)

Типичная оптимизационная задача - это создание новых изделий или технологий с более высоким качеством продукта и более низкой себестоимостью.

С простейшими задачами оптимизации мы уже встречались в курсе высшей математики при изучении методов поиска экстремумов функций (от лат. *Extremum* - крайнее).

В классическом математическом анализе задача отыскания экстремальных (т.е. либо максимальных, либо минимальных) значений некоторой функции решается на основе изучения поведения ее производных.

Классические методы поиска экстремумов с использованием производных применимы далеко не всегда. Существует очень много функций, которые не имеют производных как раз в точках максимумов или минимумов. Часто целевая функция не может быть задана ни формулой, ни уравнением, а её значения получают в результате измерений, либо численного счёта. В таких случаях применяют специальные математические методы оптимизации.

3.6. Современные методы компьютерного моделирования

В настоящее время широко применяется понятие "компьютерное моделирование", которое значительно шире традиционного понятия "моделирование на ЭВМ" и нуждается в уточнении, учитывающем сегодняшние реалии.

Начнем с термина **«компьютерная модель»**. В настоящее время под компьютерной моделью чаще всего понимают:

- условный образ объекта или некоторой системы объектов (или процессов), описанный с помощью взаимосвязанных компьютерных таблиц, блок-схем, диаграмм, графиков, рисунков, анимационных фрагментов, гипертекстов и т. д. и отображающий структуру и взаимосвязи между элементами объекта. Компьютерные модели такого вида мы будем называть **структурно-функциональными**;
- отдельную программу, совокупность программ, программный комплекс, позволяющий с помощью последовательности вычислений и графического отображения их результатов, воспроизводить (имитировать) процессы функционирования объекта, системы объектов при условии воздействия на объект различных, как правило случайных, факторов. Такие модели мы будем далее называть **имитационными моделями**.

Компьютерное моделирование - метод решения задачи анализа или синтеза сложной системы на основе использования ее компьютерной модели. Суть компьютерного моделирования заключается в получении количественных и качественных результатов по имеющейся модели. Качественные выводы, получаемые по результатам анализа, позволяют обнаружить неизвестные ранее свойства сложной системы: ее структуру, динамику развития, устойчивость, целостность и др. Количественные выводы в основном носят характер прогноза некоторых будущих или объяснения прошлых значений переменных, характеризующих систему.

Предметом компьютерного моделирования могут быть: экономическая деятельность фирмы или банка, промышленное предприятие, информационно-вычислительная сеть, технологический процесс, любой реальный объект или процесс, например процесс инфляции, и вообще - любая Сложная Система. Цели компьютерного моделирования могут быть различными, однако наиболее часто моделирование является, как уже отмечалось ранее, центральной процедурой системного анализа, причем под системным анализом мы далее понимаем совокупность методологических средств, используемых для подготовки и принятия решений экономического, организационного, социального или технического характера.

Компьютерная модель сложной системы должна по возможности отображать все основные факторы и взаимосвязи, характеризующие реальные ситуации, критерии и ограничения. Модель должна быть достаточно универсальной, чтобы по возможности описывать близкие по назначению объекты, и в то же время достаточно простой, чтобы позволить выполнить необходимые исследования с разумными затратами.

Все это говорит о том, что моделирование систем, рассматриваемое в целом, представляет собой скорее искусство, чем сформировавшуюся науку с самостоятельным набором средств отображения явлений и процессов реального мира. Поэтому исключительно сложной являются попытки классификации задач компьютерного моделирования или создания достаточно универсальных инструментальных средств компьютерного моделирования произвольных объектов. Однако если преднамеренно сузить класс рассматриваемых объектов, ограничившись, например, задачами компьютерного моделирования при системном анализе объектов экономико-организационного управления, то можно говорить о структурно-функциональном моделировании.

3.7. Структурно-функциональное моделирование

Истоки структурно-функционального моделирования следует искать в теоретических основах электрических цепей, электронике и радиотехнике, где впервые широко стали использоваться различные блок-схемы. Дальнейшее развитие структурно-функциональное моделирование получило в теории автоматического управления (ТАУ), где был развит аппарат, включающий в себя не только правила составления и преобразования, но и достаточно общую методологию анализа и синтеза структурных схем, основанную на том, что каждой математической операции над сигналами поставлен в соответствие определенный элементарный структурный блок. Хотя динамические структурно-функциональные схемы теории автоматического управления обладают широчайшими возможностями для анализа непрерывных, линейных динамических систем, описываемых дифференциальными уравнениями, они плохо подходят для описания процессов в экономико-организационных системах, где связи между отдельными блоками имеют гораздо более широкое толкование и редко могут быть сведены к некоторой функции времени (сигналу). Не очень удобны они и для описания алгоритмов и программ, для которых понятие "элементарный блок" существенно отличается от принятого в ТАУ. В частности, для составления блок-схем алгоритмов и программ, потребность в которых появилась в начале 60-х, понадобились символы, соответствующие основным операциям машинной обработки данных, их накоплению, сортировке и передаче. В результате довольно длительной разработки и последующей эволюции были созданы и нашли широкое применение государственные стандарты на составление и использование блок-схем алгоритмов и программ, вошедшие впоследствии в перечень обязательных документов Единой системы программной документации (ЕСПД). Использование стандартов на блок-схемы алгоритмов и программ весьма жестко контролировалось как Госфондом алгоритмов и программ (ГАП), так и другими "компетентными органами", причем описание любой программы и любого алгоритма должно было содержать блок-схему, даже и при отсутствии особой нужды.

Дальнейшее развитие блок-схем, связано с развитием автоматизированных систем управления производством (АСУП), появившихся в начале 70-х, в которых, в отличие от алгоритмов и программ, блок-схемы стали выполнять несколько иные функции. Основным назначением графических символов при проектировании АСУП явилось именно моделирование объекта автоматизации и процессов функционирования самой АСУП. Символика проектов АСУП включала в себя прежде всего функциональные блоки, предназначенные для отображения основных функций сбора, накопления, передачи и обработки данных. Наряду с ними в состав условных графических обозначений были включены и символы, позволяющие описывать разнообразные структуры объектов управления. На использование символов при проектировании АСУП разработаны специальные ГОСТы, регламентирующие состав, размеры и вид символов, а так же правила их использования. В целом, совокупность символов для АСУП и правил их использования образуют простейший язык структурно-функционального моделирования, применяющийся при системном анализе и проектировании автоматизированных экономико-организационных систем. Можно только сожалеть о том, что развитие подобных языков моделирования в СССР приостановилось в начале 80-х, однако в последние годы ситуация в этой области стала меняться к лучшему благодаря появлению отечественного инструментально-программного комплекса «CASE-Аналитик».

Современные методы структурно-функционального анализа и моделирования сложных систем были заложены благодаря трудам профессора Массачусетского технологического института Дугласа Росса, который впервые использовал понятие "структурный анализ" сорок лет назад, пытаясь создать алгоритмический язык АРТ, ориентированный на модульное программирование. Дальнейшее развитие идеи описания сложных объектов как иерархических, многоуровневых модульных систем с помощью относительно небольшого набора ти-

повых элементов привело к появлению SADT (Structured Analyses and Design Technique), что в дословном переводе означает "технология структурного анализа и проектирования", а по существу является методологией структурно-функционального моделирования и анализа сложных систем [10]. Со времени своего появления SADT постоянно совершенствовалась и широко использовалась для эффективного решения целого ряда проблем - таких как совершенствование управления финансами и материально-техническим снабжением крупных фирм, разработка программного обеспечения АСУ телефонными сетями, долгосрочное и стратегическое планирование деятельности фирм, проектирование вычислительных систем и сетей и др.

Центральной идеей SADT является, по определению авторов, SA-блок - универсальная единица универсальной пунктуации для неограниченного строго структурного анализа. Несмотря на такое мудреное название под таинственным SA-блоком скрывается обычный функциональный блок, характеризующийся наличием **входа, выхода, механизма и управления**. Другим фундаментальным понятием SADT является принцип иерархической декомпозиции сверху вниз, позволяющий анализировать сколь угодно сложные системы. При ближайшем рассмотрении его тоже открытием не назовешь, так как любой метод структурного анализа использует декомпозицию, которая собственно и составляет один из основных принципов познания. Оригинальным же в SADT является эффективный метод кодирования связей, основанный на использовании специальных ICOM-кодов и позволивший не только упростить процедуру моделирования, но и автоматизировать процедуры структурно-функционального анализа.

Одним из первых программных комплексов структурно-функционального анализа на основе SADT был пакет AUTOIDEF0, разработанный в рамках программы BBC США по созданию интегрированной автоматизированной системы управления производством (Integrated Computer Aided Manufacturing). В основе пакета лежит подмножество SADT, названное IDEF0. AUTOIDEF0 предназначался для облегчения процесса создания и рецензирования SADT-диаграмм и моделей для географически удаленных аэрокосмических подрядчиков. Поскольку модели часто рецензировались и исправлялись, система функционировала на диалоговых устройствах и сетях связи и включала в себя тогда еще редкие дисплеи с векторной графикой и графопостроители. Система AUTOIDEF0 предоставляла удаленным пользователям командно-ориентированную графическую среду, управляемую с помощью иерархического меню, которое облегчало работу с библиотекой диаграмм и графическими средствами. Одновременно могло создаваться, храниться, обрабатываться, публиковаться и архивироваться множество различных моделей, построенных по единой методологии средствами SADT.

Вопросы для самоконтроля

1. Какие математические модели называют дескриптивными.
2. В чём состоит различие между дескриптивными и оптимизационными моделями
3. Какие бывают физические модели
4. Что принято называть математической моделью
5. Какова классификация математических моделей
6. Что называют детерминированными математическими моделями
7. В чём отличие вероятностных моделей от детерминированных
8. Что такое имитационное математическое моделирование
9. В чём отличие компьютерного и обычного математического моделирования
10. Что называют структурно-функциональным моделированием

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 4

СОВРЕМЕННЫЕ МЕТОДЫ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ И МОДЕЛИРОВАНИЯ

4.1. Вводные замечания

Одним из необходимых условий дальнейшего развития экономической науки является применение точных методов количественного анализа, широкое использование математики. В настоящее время новейшие достижения математики и современной вычислительной техники находят все более широкое применение в экономических исследованиях и планировании. Этому способствует развитие таких разделов математики, как математическое программирование, теория игр, теория массового обслуживания, а также бурное развитие быстродействующей электронно-вычислительной техники. Уже накоплен достаточный опыт постановки и решения экономических задач с помощью математических методов. Особенно успешно развиваются методы оптимального планирования, которые и составляют сущность математического программирования.

Значительное число задач, возникающих в обществе, связано с управляемыми явлениями, т.е. с явлениями, регулируемые на основе сознательно принимаемых решений. При ограниченном объеме информации, который был доступен на ранних этапах развития общества, принималось оптимальное в некотором смысле решение на основании интуиции и опыта, а затем, с возрастанием объема информации об изучаемом явлении, - с помощью ряда прямых расчетов. Так происходило, например, создание календарных планов работы промышленных предприятий.

Совершенно иная картина возникает на современном промышленном предприятии с многосерийным и многономенклатурным производством, когда объем входной информации столь велик, что его обработка с целью принятия определенного решения невозможна без применения компьютеров. Еще большие трудности возникают в связи с задачей о принятии наилучшего решения. Проблема принятия решений в исследовании операций неразрывно связана с процессом моделирования.

Первый этап процесса моделирования состоит в построении качественной модели. Вторым этапом - построение математической модели рассматриваемой проблемы. Этот этап включает также построение целевой функции, т.е. такой числовой характеристики, большему (или меньшему) значению которой соответствует лучшая ситуация с точки зрения принимающего решения. В результате этих двух этапов формируется соответствующая математическая задача.

Третий этап - исследование влияния переменных на значение целевой функции. Этот этап предусматривает владение математическим аппаратом для решения математических задач, возникающих на втором этапе процесса принятия решения.

Четвертый этап - сопоставление результатов вычислений, полученных на третьем этапе, с моделируемым объектом, т.е. экспертная проверка результатов (критерий практики). Таким образом, на этом этапе устанавливается степень адекватности модели и моделируемого объекта в пределах точности исходной информации.

Широкий класс задач управления составляют такие экстремальные задачи, в математических моделях которых условия на переменные задаются равенствами и неравенствами. Теория и методы решения этих задач как раз и составляют содержание математического программирования.

4.2. Линейное математическое программирование

В настоящее время линейное программирование является одним из наиболее употребительных аппаратов математической теории оптимального принятия решений. Для решения

задач линейного программирования разработано сложное программное обеспечение, дающее возможность эффективно и надежно решать практические задачи больших объемов. Владение аппаратом линейного программирования необходимо каждому специалисту в области прикладной математики.

Линейное программирование – это наука о методах исследования и отыскания наибольших и наименьших значений линейной функции, на неизвестные которой наложены линейные ограничения. Таким образом, задачи линейного программирования относятся к задачам на условный экстремум функции. По типу решаемых задач методы разделяются на универсальные и специальные. С помощью универсальных методов могут решаться любые задачи линейного программирования (ЗЛП). Специальные методы учитывают особенности модели задачи, ее целевой функции и системы ограничений.

Особенностью задач линейного программирования является то, что экстремума целевая функция достигает на границе области допустимых решений. Классические же методы дифференциального исчисления связаны с нахождением экстремумов функции во внутренней точке области допустимых значений. Отсюда — необходимость разработки новых методов.

Линейное программирование представляет собой наиболее часто используемый метод оптимизации. К числу задач линейного программирования можно отнести задачи:

1. рационального использования сырья и материалов;
2. задачи оптимального раскроя;
3. оптимизации производственной программы предприятий;
4. оптимального размещения и концентрации производства;
5. составления оптимального плана перевозок, работы транспорта (транспортные задачи);
6. управления производственными запасами;
7. многие другие, принадлежащие сфере оптимального планирования.

Линейное программирование является одной из основных частей раздела современной математики, который получил название математического программирования. В общей постановке, задачи этого раздела выглядят следующим образом.

Требуется найти такие неотрицательные $X_1, X_2, X_3, \dots, X_N$, которые обеспечивают максимум или минимум целевой функции (формула 4.1), которые удовлетворяют системе ограничений (формула 4.2) и не противоречат условиям неотрицательности:

$$F(X) = C_1 X_1 + C_1 X_2 + C_1 X_3 + \dots + C_1 X_{1N} \quad (4.1)$$

$$\begin{aligned}
 & A_{11}\mathbf{X}_1 + A_{12}\mathbf{X}_2 + A_{13}\mathbf{X}_3 + \dots + A_{1N}\mathbf{X}_1 \quad \{\leq =\geq\} B_1 \\
 & A_{21}\mathbf{X}_1 + A_{22}\mathbf{X}_2 + A_{23}\mathbf{X}_3 + \dots + A_{2N}\mathbf{X}_1 \quad \{\leq =\geq\} B_2 \\
 & A_{11}\mathbf{X}_1 + A_{12}\mathbf{X}_2 + A_{13}\mathbf{X}_3 + \dots + A_{1N}\mathbf{X}_1 \quad \{\leq =\geq\} B_3 \\
 & \\
 & A_{M1}\mathbf{X}_1 + A_{M2}\mathbf{X}_2 + A_{M3}\mathbf{X}_3 + \dots + A_{MN}\mathbf{X}_1 \quad \{\leq =\geq\} B_M
 \end{aligned} \tag{4.2}$$

В зависимости от вида функции различают разделы математического программирования: квадратичное программирование, выпуклое программирование, целочисленное программирование и т.д. Линейное программирование характеризуется тем, что функция является линейной функцией переменных.

Задачу на минимум можно решать как задачу на максимум. Достаточно знаки целевой функции поменять на противоположные.

Целевая функция задачи линейного программирования достигает своего экстремума (минимума или максимума) в вершине допустимой области. Если целевая функция достигает экстремального значения более чем на одной вершине, то она достигает того же значения в любой точке, являющейся выпуклой комбинацией этих вершин (альтернативный оптимум).

Эта теорема имеет важнейшее значение, так как она указывает путь решения задачи линейного программирования. Совсем не надо перебирать все точки допустимой области. Дос-

таточно перебрать вершины допустимой области, а ведь их конечное число. Кроме того, не нужно перебирать все вершины, можно этот перебор существенно сократить.

Любой набор чисел, удовлетворяющий ограничениям задачи, называют планом, а множество всех планов допустимой областью. Тот план, который доставляет экстремум (минимум или максимум) целевой функции, называют оптимальным планом или просто решением задачи линейного программирования.

4.3. Методы решения задач линейного программирования

Задачи линейного программирования решаются несколькими методами:

1. графический метод;
2. симплексный метод;
3. двойственность в ЛП;
4. двойственный симплексный метод.

Задачи линейного программирования с двумя переменными всегда можно решить графически. Однако уже в трехмерном пространстве такое решение усложняется, а в пространствах, размерность которых больше трех, графическое решение невозможно.

Графический метод довольно прост и нагляден. Он основан на геометрическом представлении допустимых решений задачи. Каждое из неравенств задачи ЛП определяет на координатной плоскости некоторую полуплоскость, а система неравенств в целом – пересечение соответствующих плоскостей. Множество точек пересечения данных полуплоскостей называется областью допустимых решений (ОДР). ОДР всегда представляет собой выпуклую фигуру, т.е. обладающую следующим свойством: если две точки А и В принадлежат этой фигуре, то и весь отрезок АВ принадлежит ей. ОДР графически может быть представлен выпуклым многоугольником, неограниченным выпуклой многоугольной областью, отрезком, лучом и т.д. В случае несовместности системы ограничений задачи ОДР является пустым множеством.

При поиске оптимального решения задач линейного программирования возможны следующие ситуации: существует единственное решение задачи, существует бесконечное множество решений (альтернативный оптимум); ЦФ не ограничена; область допустимых решений – единственная точка; задача не имеет решений.

Любая задача линейного программирования, независимо от вида записи, может быть приведена к стандартной и канонической форме и решена симплексным методом, который в определенном смысле является универсальным методом ЛП. Алгоритм симплекс-метода носит итерационный характер.

Симплекс-метод позволяет переходить от одного допустимого базисного решения к другому, причем так, что значения целевой функции непрерывно возрастают. Алгоритмы симплекс-метода позволяют также установить, является ли задача ЛП разрешимой.

Переход от одного базиса к другому позволяет находить решения почти всех задач ЛП. Определив все крайние точки, можно вычислить значения целевой функции и найти оптимальное решение. Однако для больших значений m и n это практически невозможно.

Алгоритм решения задачи ЛП табличным симплексом-методом состоит из следующих этапов:

1. рассчитывают и заполняют начальную симплекс-таблицу с допустимым единичным базисом, включая индексную строку;
2. находят разрешающий столбец;
3. находят разрешающую строку;
4. рассчитывают методом Жордано-Гаусса все параметры матрицы;
5. анализируют полученные данные в индексной строке.

Таблицы симплекс-метода необходимо строить до тех пор, пока не будет получен оптимальный план. План будет считаться оптимальным, если в последней индексной строке симплекс-таблицы будут только нули и положительные числа.

При построении симплексного метода предполагалось, что все опорные планы невырожденные, что обеспечивало получение оптимального плана за конечное количество шагов. В случае вырожденного плана вычисления производят аналогично, но в этом случае возможен возврат к старому базису, что приводит к так называемому заикливанию.

Метод искусственного базиса применяется при наличии в ограничении знаков “равно”, “больше либо равно”, “меньше либо равно” и является модификацией табличного метода. Решение системы производится путём ввода искусственных переменных со знаком, зависящим от типа оптимума, т.е. для исключения из базиса этих переменных последние вводятся в целевую функцию с большими отрицательными коэффициентами m , а в задачи минимизации - с положительными m . Таким образом, из исходной получается новая m - задача.

Если в оптимальном решении m - задачи нет искусственных переменных, это решение есть оптимальное решение исходной задачи. Если же в оптимальном решении m - задачи хоть одна из искусственных переменных будет отлична от нуля, то система ограничений исходной задачи несовместна и исходная задача неразрешима.

В основу модифицированного симплекс – метода положены такие особенности линейной алгебры, которые позволяют в ходе решения задачи работать с частью матрицы ограничений. Иногда метод называют методом обратной матрицы.

В процессе работы алгоритма происходит спонтанное обращение матрицы ограничений по частям, соответствующим текущим базисным векторам. Указанная способность делает весьма привлекательной машинную реализацию вычислений вследствие экономии памяти под промежуточные переменные и значительного сокращения времени счёта. Метод эффективен для ситуаций, когда число переменных n значительно превышает число ограничений m .

В целом, метод отражает традиционные черты общего подхода к решению задач линейного программирования, включающего в себя канонизацию условий задачи, расчёт симплекс-разностей, проверку условий оптимальности, принятие решений о коррекции базиса и исключение Жордана-Гаусса.

Особенности заключаются в наличии двух таблиц - основной и вспомогательной, порядке их заполнения и некоторой специфичности расчётных формул.

4.4. Двойственная задача линейного программирования

Каждой задаче линейного программирования можно определенным образом сопоставить некоторую другую задачу, называемую двойственной или сопряженной по отношению к исходной или прямой задаче. Сопоставляя формы записи прямой и двойственной задач, можно установить между ними следующие взаимосвязи:

1. если прямая задача является задачей максимизации, то двойственная будет задачей минимизации, и наоборот;
2. коэффициенты целевой функции прямой задачи становятся свободными членами ограничений двойственной задачи;
3. свободные члены ограничений прямой задачи становятся коэффициентами целевой функции двойственной задачи;
4. матрица ограничений двойственной задачи получается путем транспонирования матрицы ограничений прямой задачи;
5. знаки неравенств в ограничениях изменяются на противоположные;
6. число ограничений прямой задачи равно числу переменных двойственной задачи, и наоборот.

Если из пары двойственных задач одна обладает оптимальным планом, то и другая имеет решение, причем для экстремальных значений линейных функций выполняется определенное соотношение. Если линейная функция одной из задач не ограничена, то другая не имеет решения.

Если прямая (а значит, и двойственная) задача разрешима, то в каждой паре двойственных условий одно является свободным, а другое закрепленным. Любое из условий называется свободным, если оно выполняется как строгое неравенство хотя бы для одного оптимального вектора. Условие называется закрепленным, если оно выполняется как равенство для всех оптимальных векторов.

Двойственную задачу выгоднее решать, чем прямую, если в прямой задаче при малом количестве переменных имеется большое количество ограничений.

4.5. Двойственный симплекс-метод

Симплексный метод позволяет наряду с получением решения прямой задачи получать и решение двойственной задачи. Этот результат и лежит в основе двойственного симплексного метода решения задачи. Суть метода состоит в таком последовательном переборе угловых точек допустимого множества Q_0 двойственной задачи, при котором значение целевой функции возрастает, т. е. в применении симплексного метода к решению двойственной задачи. Будем предполагать, что задача невырождена, т. е. каждой угловой точке множества Q_0 соответствует квадратная невырожденная система уравнений размерности m , матрицу которую и называют двойственным базисом прямой задачи. Вместе с тем двойственный симплекс-метод можно применять при решении задачи линейного программирования, свободные члены системы уравнений которой могут быть любыми числами (при решении задачи симплексным методом эти числа предполагались неотрицательными).

Отыскание решения задачи двойственным симплекс-методом включает в себя следующие этапы:

1. Находят псевдоплан задачи.

2. Проверяют этот псевдоплан на оптимальность. Если псевдоплан оптимален, то найдено решение задачи. В противном случае либо устанавливают неразрешимость задачи, либо переходят к новому псевдоплану.

3. Выбирают разрешающую строку с помощью определения наибольшего по абсолютной величине отрицательного числа столбца вектора P_0 и разрешающий столбец с помощью нахождения наименьшего по абсолютной величине отношения элементов $(m+1)$ -и строки к соответствующим отрицательным элементам разрешающей строки.

4. Находят новый псевдоплан и повторяют все действия начиная со второго этапа.

Двойственный симплексный метод называют также методом последовательного уточнения оценок, поскольку угловые точки задачи, возникающие при итерациях, можно рассматривать как приближенные значения точной оценки y^* , т. е. как приближенные оценки влияния условий задачи на величину минимума целевой функции.

4.6. Целочисленное математическое программирование

Значительная часть экономических задач, относящихся к задачам линейного программирования, требует целочисленного решения. К ним относятся задачи, у которых переменные величины означают количество единиц неделимой продукции, например распределение производственных заданий между предприятиями, раскрой материалов, загрузка оборудования, распределение судов по линиям, самолетов по рейсам, а также задачи по производству неделимой продукции. Если единица составляет малую часть всего объема производства, то оптимальное решение находят обычным симплексным методом, округляя его до целых единиц, исходя из смысла задачи. В противном случае округление может привести к решению, далекому от оптимального целочисленного решения.

Задача целочисленного программирования формулируется так же, как и задача линейного программирования, но включается дополнительное требование, состоящее в том, что значения переменных, составляющих оптимальное решение, должны быть целыми неотрицательными числами.

Метод решения таких задач, предложенный Гомори, основан на симплексном методе и состоит в следующем. Симплексным методом находится оптимальный план задачи без учета условия целочисленности. Если оптимальный план целочисленный, то вычисления заканчивают; если же оптимальный план содержит хотя бы одну дробную компоненту X_i , то накладывают дополнительное ограничение, учитывающее целочисленность компонент плана, и вычисления симплексным методом продолжают до тех пор, пока либо будет найден целочисленный оптимальный план, либо доказано, что задача не имеет целочисленных оптимальных планов.

Особенно широкое распространение линейное программирование получило в экономике, так как исследование зависимостей между величинами, встречающимися во многих экономических задачах, приводит к линейной функции с линейными ограничениями, наложенными на неизвестные.

4.7. Области применения и ограничения линейного программирования

Особенно широкое применение методы и модели линейного программирования получили при решении задач экономии ресурсов (выбор ресурсосберегающих технологий, составление смесей, раскрой материалов, производственно-транспортных и других задач).

Рассмотрим постановку задачи о наилучшем использовании ресурсов. Пусть некоторая производственная единица (цех, завод, объединение и т. д.), исходя из конъюнктуры рынка, технических или технологических возможностей и имеющихся ресурсов, может выпускать n различных видов продукции (товаров), известных под номерами, обозначаемыми индексом j . Товары будем обозначать X_j . Предприятие при производстве этих видов продукции должно ограничиваться имеющимися видами ресурсов, технологий, других производственных факторов (сырья, полуфабрикатов, рабочей силы, оборудования, электроэнергии и т. д.). Все эти виды ограничивающих факторов называют ингредиентами. Пусть их число равно m ; припишем им индекс i . Они ограничены, и их количества равны соответственно условных единиц. Таким образом, R_i – вектор ресурсов. Известна экономическая выгода (мера полезности) производства продукции каждого вида, исчисляемая, скажем, по отпускной цене товара, его прибыльности, издержкам производства, степени удовлетворения потребностей и т. д. Примем в качестве такой меры, например, цену реализации, т. е. C_j – вектор цен. Известны также технологические коэффициенты a_{ij} , которые указывают, сколько единиц i -го ресурса требуется для производства единицы продукции j -го вида. Матрицу коэффициентов называют технологической и обозначают буквой A . Имеем X_j . Обозначим через Z план производства, показывающий, какие виды товаров нужно производить и в каких количествах, чтобы обеспечить предприятию максимум объема реализации при имеющихся ресурсах. Так как C_j – цена реализации единицы j -й продукции, цена реализованных единиц будет равна $C_j X_j$, а общий объем реализации примет вид $Z = \sum C_j X_j$. Это — целевая функция, которую нужно максимизировать.

Так как a_{ij} – расход i -го ресурса на производство единиц j -й продукции, то, просуммировав расход i -го ресурса на выпуск всех n видов продукции, получим общий расход этого ресурса, который не должен превосходить R_i единиц.

Чтобы искомым план был реализован, наряду с ограничениями на ресурсы нужно наложить условие неотрицательности на объемы выпуска продукции.

В модель задачи о наилучшем использовании ресурсов входят: целевая функция, система ограничений и условия неотрицательности

Так как переменные X_j входят в функцию Z и систему ограничений только в первой степени, а показатели a_{ij} являются постоянными в планируемый период, то это – задача линейного программирования.

В различных отраслях народного хозяйства возникает проблема составления таких рабочих смесей на основе исходных материалов, которые обеспечивали бы получение конечного продукта, обладающего определенными свойствами. К этой группе задач относятся задачи о выборе диеты, составлении кормового рациона в животноводстве, шихт в металлур-

гии, горючих и смазочных смесей в нефтеперерабатывающей промышленности, смесей для получения бетона в строительстве и т. д.. Высокий уровень затрат на исходные сырьевые материалы и необходимость повышения эффективности производства выдвигает на первый план следующую задачу: получить продукцию с заданными свойствами при наименьших затратах на исходные сырьевые материалы.

Сущность задачи об оптимальном раскрое состоит в разработке таких технологически допустимых планов раскроя, при которых получается необходимый комплект заготовок, а отходы (по длине, площади, объему, массе или стоимости) сводятся к минимуму. Более сложные постановки ведут к задачам целочисленного программирования.

Общая постановка транспортной задачи состоит в определении оптимального плана перевозок некоторого однородного груза из m пунктов отправления в n пунктов назначения. При этом в качестве критерия оптимальности обычно берется либо минимальная стоимость перевозок всего груза, либо минимальное время его доставки. Рассмотрим транспортную задачу, в качестве критерия оптимальности которой взята минимальная стоимость перевозок всего груза. Обозначим через c_{ij} тарифы перевозки единицы груза из i -го пункта отправления в j -й пункт назначения, через a_i – запасы груза в i -м пункте отправления, через b_j – потребности в грузе в j -м пункте назначения, а через x_{ij} – количество единиц груза, перевозимого из i -го пункта отправления в j -й пункт назначения.

Тогда математическая постановка задачи состоит в определении минимального значения целевой функции при определенных ограничениях и условиях неотрицательности.

Таким образом, обеспечивается доставка необходимого количества груза в каждый из пунктов назначения, вывоз имеющегося груза из всех пунктов отправления, а также исключаются обратные перевозки. Всякое неотрицательное решение систем линейных уравнений называется планом транспортной задачи. План, при котором целевая функция принимает свое минимальное значение, называется оптимальным планом транспортной задачи. Если в опорном плане число отличных от нуля компонент равно в точности $n+m-1$, то план является невырожденным, а если меньше – то вырожденным.

Если общая потребность в грузе в пунктах назначения равна запасу груза в пунктах отправления, то модель такой транспортной задачи называется закрытой. Если же указанное условие не выполняется, то модель транспортной задачи называется открытой.

В случае превышения запаса над потребностью, вводится фиктивный $(n+1)$ -й пункт назначения с потребностью и соответствующие тарифы считаются равными нулю. Аналогично, в случае, если потребности превышают количество запасов, также вводится фиктивный $(m+1)$ -й пункт отправления с запасом груза и тарифы полагаются равными нулю. Этим задача сводится к обычной транспортной задаче, из оптимального плана которой получается оптимальный план исходной задачи.

4.8. Методы построения опорных планов

Как и для всякой задачи линейного программирования, оптимальный план транспортной задачи является и опорным планом. Опорный план является допустимым решением ТЗ и используется в качестве начального базисного решения при нахождении оптимального решения методом потенциалов. Существует четыре метода нахождения опорных планов:

1. метод северо-западного угла;
2. метод минимального элемента;
3. метод двойного предпочтения;
4. метод штрафов (Фогеля).

"Качество" опорных планов, полученных этими методами, различается: в общем случае метод Фогеля дает наилучшее решение (зачастую оптимальное), а метод северо-западного угла – наихудшее.

Все существующие методы нахождения опорных планов отличаются только способом выбора клетки для заполнения. Само заполнение происходит одинаково независимо от ис-

пользуемого метода. Следует помнить, что перед нахождением опорного плана транспортная задача должна быть сбалансирована.

В методе северо-западного угла из всех не вычеркнутых клеток выбирается самая левая и верхняя (северо-западная) клетка. Другими словами, на каждом шаге выбирается первая из оставшихся не вычеркнутых строк и первый из оставшихся не вычеркнутых столбцов.

Для того чтобы заполнить клетку (i, j) , необходимо сравнить текущий запас товара в рассматриваемой i -й строке с текущей потребностью в рассматриваемом j -м столбце. Нахождение опорного плана продолжается до тех пор, пока не будут вычеркнуты все строки и столбцы.

В методе минимального элемента первой клеткой выбирают клетку с наименьшей суммой доставки и заполняют ее максимально возможным грузом.

Если таблица стоимостей велика, то перебор всех элементов затруднителен. В этом случае используют метод двойного предпочтения, суть которого заключается в следующем: в каждой строке и каждом столбце отмечают «V» наименьшую стоимость, а затем клетки с двойным символом «VV» заполняют с учетом наименьшей стоимости. Затем распределяют перевозки по клеткам, отмеченным знаком «V». В оставшейся части таблицы перевозки распределяют по наименьшей стоимости.

На каждом шаге метода Фогеля для каждой i -й строки вычисляются штрафы, как разность между двумя наименьшими тарифами строки. Таким же образом вычисляются штрафы для каждого j -го столбца. После чего выбирается максимальный штраф из всех штрафов строк и столбцов. В строке или столбце, соответствующем выбранному штрафу, для заполнения выбирается не вычеркнутая клетка с минимальным тарифом. Если существует несколько одинаковых по величине максимальных штрафов в матрице, то в соответствующих строках или столбцах выбирается одна не вычеркнутая клетка с минимальным тарифом.

Если клеток с минимальным тарифом также несколько, то из них выбирается клетка (i, j) с максимальным суммарным штрафом, т.е. суммой штрафов по i -й строке и j -му столбцу.

Если план транспортной задачи является оптимальным, то ему соответствует система из $m+n$ чисел U_i и V_j , удовлетворяющих условиям: $U_i + V_j = C_{ij}$ для занятых клеток и $U_i + V_j \leq C_{ij}$ в свободных клетках. Числа U_i и V_j называются потенциалами соответственно поставщиков и потребителей. При решении одному неизвестному потенциалу придается произвольное значение.

Задачи математического программирования находят применение в различных областях человеческой деятельности, где необходим выбор одного из возможных образов действий (программ действий), например, при решении проблем управления и планирования производственных процессов, в проектировании и перспективном планировании, в военном деле и т.д.

Вопросы для самоконтроля

1. Определение линейного программирования и моделирования.
2. Основные этапы разработки модели.
3. Основные приёмы моделирования.
4. Какие виды задач линейного программирования вы знаете?
5. Открытые и закрытые транспортные задачи.

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

ВЕРОЯТНОСТНЫЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

5.1. Основные определения

Вероятностными называют такие процессы, в которых соответствие входных и выходных переменных носит случайный характер и описывается распределением вероятностей.

Для математического моделирования таких процессов используются методы теории вероятностей и математической статистики.

В литературе в качестве синонимов вероятностных процессов используют иногда термины стохастические, статистические и эмпирические. «Эмпирические» (от греческого *empeiria* – опыт) потому, что параметры этих процессов приходится оценивать путем статистической обработки результатов опыта, эксперимента.

С вероятностными процессами мы имеем дело практически каждый день. В частности, подходя утром к остановке трамвая или троллейбуса, мы не знаем заранее, сколько времени придется ждать вагон.

Можно только предполагать, что трамвай появится в течение 5-15 минут, но сколько конкретно: 6, 9 или 14 минут – предугадать невозможно. Расписание движения существует, но на фактический интервал движения оказывает влияние много факторов, которые нельзя учесть заранее. Например, на одной остановке пассажиров было мало, и время стоянки сократилось (интервал между трамваями по времени сократился). На перекрестке образовался затор – интервал увеличился. Из-за сложности и неопределенности многих внешних воздействий интервал движения трамваев оказывается величиной случайной, подчиняющейся вероятностным законам.

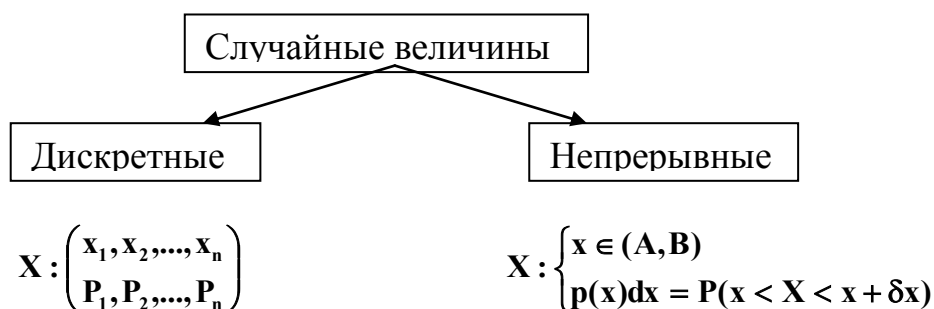
Случайной величиной называется величина, которая в результате опыта (эксперимента) может принять одно возможное, но заранее неизвестное значение.

Говоря о том, что интервал движения является величиной случайной, мы утверждаем, что не можем указать, какое значение он примет в данный конкретный момент, например, 28 апреля 1994 года в 1235. Однако на основании своего жизненного опыта мы знаем, какие значения он может принимать вообще (от 1 минуты до 1 часа!) и знаем вероятности тех или иных значений. Эти знания позволяют нам свободно ориентироваться в обстановке на общественном транспорте. Так, если трамвая нет более 15-20 минут, то значит, что где-то произошла авария и лучше идти пешком или бежать на остановку троллейбуса (особенно утром). Итак, чтобы задать случайную величину, необходимо указать, какие значения она может принимать и какова вероятность появления этих значений.

Случайные величины принято разделять на дискретные и непрерывные.

Случайная величина называется дискретной (от лат. *Discretus* – прерывистый, состоящий из отдельных частей), если она может принимать только конечные или счетное множество значений.

Счетное множество – это бесконечное множество чисел, содержащее столько чисел, сколько и в бесконечном ряду натуральных чисел: 0, 1, 2, 3, 4, 5, 6, 7, ...



С каждым из возможных значений дискретной случайной величины $x_i, i=1,2,3,\dots,N$, связывают некоторое число $P(x_i)$, называемое вероятностью появления этого значения x . Число $P(x_i)$ характеризует среднее число (частоту) появления данного значения при многократных повторениях опыта или эксперимента (измерения). Например, если целый месяц каждое утро наблюдать интервалы движения трамваев, легко установить, что чаще всего трамвай появляется с интервалом 5-10 минут. Следовательно, вероятность этих значений наибольшая. А частота прихода трамвая через 1 секунду или через 1 час – очень мала и соответствующая вероятность – тоже!

Любая дискретная случайная величина определяется таблицей:

$$X \sim \begin{pmatrix} x_1, x_2, x_3, \dots, x_n \\ p_1, p_2, p_3, \dots, p_n \end{pmatrix}, \text{ где } x_1, x_2, \dots, x_n - \text{возможные значения величины } x, \text{ а } p_1, p_2,$$

$p_3, \dots, p_n$ – соответствующие им вероятности. Эта таблица называется распределением случайной величины X .

Значения случайной величины $x_i, i=1,2,3,\dots,N$ могут быть (в общем случае) любыми, но вероятности $p_i, i=1,2,3,\dots,N$ должны удовлетворять двум условиям:

Все p_i неотрицательны:

$$p_i \geq 0$$

Сумма всех p_i равна 1:

$$p_1 + p_2 + p_3 + \dots + p_n = 1.$$

Условие (2) означает, что случайная величина X в каждом опыте принимает какое-то одно конкретное значение из ее возможных значений $x_1, x_2, \dots, x_n$.

5.2 Числовые характеристики случайных величин

Задание распределения случайной величины полностью характеризует ее поведение. Однако на практике часто не требуется такого детального описания. Так, для пассажира важно не полное распределение интервалов движения трамвая по часам и минутам, а только наиболее вероятные значения интервала (чтобы знать, сколько времени имеет смысл ждать), а также каков разумный разброс времени ожидания. Такую информацию дают числовые характеристики распределения случайной величины, называемые в теории вероятности математическим ожиданием, а в математической статистике – выборочным средним.

Математическим ожиданием называется число:

$$MX = \sum_{i=1}^n x_i p_i.$$

Оно характеризует наиболее вероятные значения случайной величины X .

Разброс отдельных значений x_i относительно математического ожидания характеризует еще один важный числовой параметр распределения: дисперсия (от лат. рассеяние, рассыпание):

$$DX = M(X - MX)^2$$

Дисперсия – это математическое ожидание квадрата отклонения случайной величины от ее математического ожидания (т.е. среднего значения).

В математической статистике среднее значение случайной величины определяется как:

$$\bar{X} = \langle X \rangle = X_{cp} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \approx MX,$$

где $x_i, i=1,2,3,\dots,N$ – случайная величина.

Дисперсия:

$$D = \frac{(x_1 - X_{cp})^2 + (x_2 - X_{cp})^2 + (x_3 - X_{cp})^2 + \dots + (x_n - X_{cp})^2}{n - 1}$$

В практике обработки экспериментальных данных наряду с дисперсией используется еще и среднее квадратическое отклонение:

$$\sigma = \sqrt{D}.$$

5.3. Вычисление числовых характеристик случайных величин

Формулы для вычисления средних значений, дисперсии и среднего квадратического отклонения просты, и их программирование не вызывает затруднений даже у обучающихся 1 курса.

Проблемы возникают при вводе данных. Здесь возможны две ситуации:

Число данных (т.е. число чисел $x_1, x_2, \dots, x_n$) известно до начала работы с программой.

Число данных N произвольно и заранее неизвестно.

Для каждого из этих случаев целесообразно использовать свой способ организации программы на ПК.

```
Program AVERAGE1; {вычисление среднего для фиксированного числа слагаемых}
Const N=10; {N –число слагаемых }
Var X, Sum:real; {X – текущее слагаемое, Sum – суммаX}
I:integer; {I – число повторов}
Begin
  Writeln ('Введите через пробел',N:4,'-элементный список');
  Sum:=0; I:=0;
  While I<N DO
    Begin
      Read (X);
      Sum:=Sum+X;
      I:=I+1
    End;
  Writeln ('Среднее значение =',Sum/N:6:4)
End.
```

NB: Здесь - чтобы получить результаты или просто вернуться в окно редактирования, необходимо ввести все 10 чисел!

Часто число слагаемых при вычислении среднего заранее неизвестно. Тогда для окончания ввода данных используют некоторое специальное значение, называемое признаком конца ввода или «стоп-кодом». Это значение заведомо не может встретиться в потоке входных данных. Например, если вычисляется средний балл по результатам всех модулей, то среди отметок никак не может появиться отрицательное число (например, -1). Следовательно, ввод такого, заведомо нереального, слагаемого можно использовать для прекращения ввода данных.

```
Program AVERAGE2; {вычисление среднего произвольного числа слагаемых с прекращением ввода с помощью стоп-кода}
Const SC=-1; {SC – «стоп-код»}
Var X, Sum:real; {X – текущее слагаемое, Sum – суммаX}
I:integer; {I – число повторов}
Begin
  Writeln ('Введите через пробел числа');
  Writeln ('Для окончания ввода наберите -1 и нажмите ENTER');
  Sum:=0; I:=0; X:=0;
  While X<>SC DO
```

```

Begin
  Sum:=Sum+X;
  I:=I+1;
  Read (X)
End;
Writeln ('Среднее значение =',Sum/(I-1):6:4)
End.

```

Программы AVERAGE1 и AVERAGE2 легко дополнить фрагментами для вычисления дисперсии и среднего квадратического отклонения. Эти программы удобно использовать для «моментального» вычисления средних значений, когда не надо запоминать значения слагаемых x_i , $i=1,2,3,\dots,N$. Если же необходимо сохранить величины x_i , $i=1,2,3,\dots,N$ для последующей обработки, то целесообразно использовать массивы.

```

Program VARIANCE;
{вычисление среднего, дисперсии и среднего квадратического отклонения}
Const N=10;
Var Sum,VR:real;
I:integer;
A:Array[1..N] of real;
Begin
  Writeln ('Введите через пробел',N:3,'-элементный список');
  For I:=1 TO N DO
    Read (A[I]);
  Sum:=0;
  For I:=1 TO N DO
    Sum:=Sum+A[I];
    Writeln ('Среднее значение=',Sum/N:6:4);
    VR:=0;
  For I:=1 TO N DO
    VR:=VR+(A[I]-Sum/N)*(A[I]-Sum/N)/(N-1);
  Writeln('Дисперсия равна', VR:6:4);
  Writeln ('Среднее квадратическое отклонение=',SQRT(VR):6:4)
End.

```

Контрольный вариант: 1 2 3 4 5 6 7 8 9 10
 Среднее = 5.5000
 Дисперсия = 9.1667
 Среднее кв.отклонение = 3.02765

5.4 Непрерывные случайные величины

Наряду с дискретными часто приходится иметь дело со случайными величинами, которые могут принимать любые, а не только отдельные (дискретные) значения. Например, интервал времени между прибытием на остановку очередного трамвая – это непрерывная случайная величина. Она может принимать значения и в 1 минуту, и в 5 минут 47 секунд, и в 1 час 3 минуты 15 секунд.

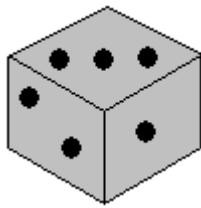
С инженерной практике чаще приходится иметь дело именно с непрерывными случайными величинами, особенно при обработке результатов самых различных измерений.

Случайная величина называется непрерывной, если она может принимать любое значения в некотором интервале $[a,b]$.

Мы уже знаем, что для задания дискретной случайной величины X необходимо задать ее распределение, т.е. таблицу возможных значений x_i и соответствующих вероятностей p_i .

Классический пример: кость. Это правильный кубик из однородного материала, на гранях которого нанесены цветные точки: одна, две, три, ..., шесть.

Распределение этой случайной (дискретной) величины – это таблица:



$$X: \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

←	Число точек на верхней грани
←	Вероятности их выпадения

Для непрерывной случайной величины такую таблицу задать нельзя, т.к. возможных значений x для непрерывной случайной величины бесконечно много. Дело в том, что даже на самом маленьком интервале числовой оси содержится («вмещается») бесконечно много чисел!

Поэтому для определения непрерывной случайной величины необходимо задать интервал (a, b) , содержащий все возможные значения величины, и закон распределения вероятностей этой случайной величины. Законом распределения вероятностей называют правило, позволяющее определять вероятность попадания случайной величины в любую заданную область ее значений.

Обычно в качестве закона распределения непрерывной случайной величины используют ту или иную известную или хорошо изученную функцию (например, экспоненту, логарифм или гипергеометрическую «гамма»-функцию).

Рассмотрим свойства распределений непрерывных случайных величин более подробно.

Закон распределения непрерывной случайной величины X может быть задан с помощью функции $P(x)$, такой, что произведение $p(x)dx$ есть вероятность попадания случайной величины X в интервал $(x, x+dx)$:

$$p(x) dx = P(x \leq X < x+dx)$$

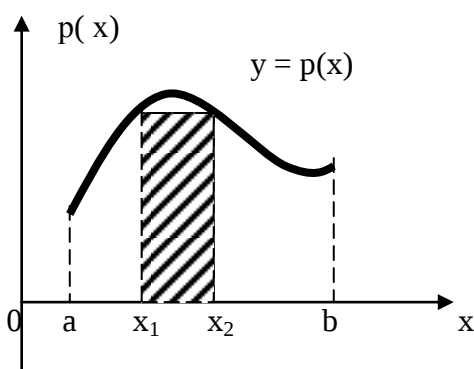
Здесь P – вероятность события, а в скобках – в чем состоит это случайное событие.

Функция $p(x)$ называется плотностью распределения вероятности или просто плотностью распределения. Плотность распределения вероятности обладает свойствами, аналогичными свойствам вероятностей дискретных случайных величин:

1. $p(x) \geq 0$ – плотность неотрицательна.
2. $\int_a^b p(x)dx = 1$ – это так называемое условие нормировки; оно означает, что с вероятностью, равной 1, случайная величина X должна принять какое-то конкретное значение из интервала (a, b) .

Зная плотность распределения, легко найти вероятность попадания случайной величины в любой заданный интервал (x_1, x_2) внутри интервала (a, b) :

$$P(x_1 \leq X < x_2) = \int_{x_1}^{x_2} p(x)dx$$



Как и для дискретных величин, так и для непрерывных, важную роль играют такие числовые характеристики, как математическое ожидание (среднее значение) и дисперсия.

Математическое ожидание:

$$MX = \int_a^b x \cdot p(x) dx = \mu$$

Оно характеризует наиболее вероятные значения, которые может принимать непрерывная случайная величина X .

Дисперсия:

$$DX = M(X - MX)^2 = \int_a^b (x - \mu)^2 P(x) dx$$

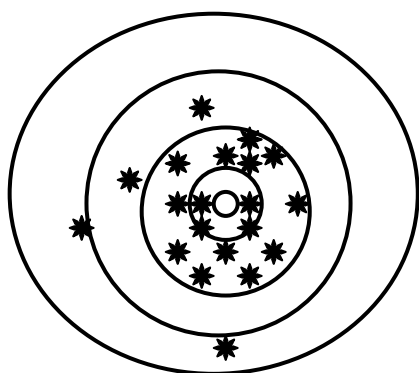
$$DX = M(X^2) - (MX)^2 = \int_a^b x^2 P(x) dx - \mu^2$$

Среднее квадратическое отклонение:

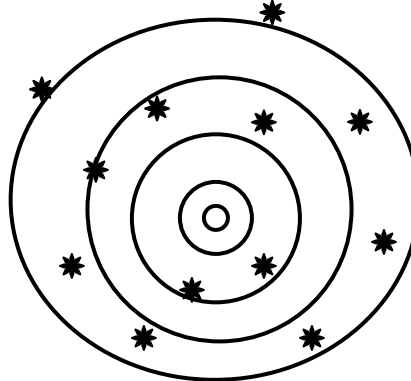
$$\sigma = \sqrt{DX}$$

Дисперсия и среднее квадратическое отклонение показывают, насколько далеко разбросаны значения случайной величины относительно ее среднего значения.

Очень наглядно можно наблюдать дисперсию в стрелковом тире.



Снайпер
(малая дисперсия)



Начинающий стрелок
(большая дисперсия)

5.5 Модельные распределения вероятностей

Законы распределения непрерывных случайных величин весьма разнообразны и определяются физической природой этих величин. Например, закон распределения роста взрослых людей явно отличается от закона распределения выпадений того или иного числа цветных точек при бросании игрального кубика. Так, очевидно, что вероятность встретить на улице очень высокого (более 2 метров) или очень маленького (менее 0,5 метра) взрослого человека, как правило, мала, а вероятность выпадения любого числа цветных точек на верхней грани игрального кубика всегда постоянна и равна $1/6$.

Строго говоря, точные законы распределения реальных процессов неизвестны, а приближенные могут быть найдены только из эксперимента. Это всегда непростая задача.

На практике при построении вероятностных моделей различных природных и технологических процессов широко используют стандартные, хорошо изученные законы распределения случайных величин. Их часто называют модельными распределениями.

Модельными законами распределения вероятностей называют такие теоретические функции (кривые), которые используются в качестве математических моделей распределений вероятностей реальных процессов.

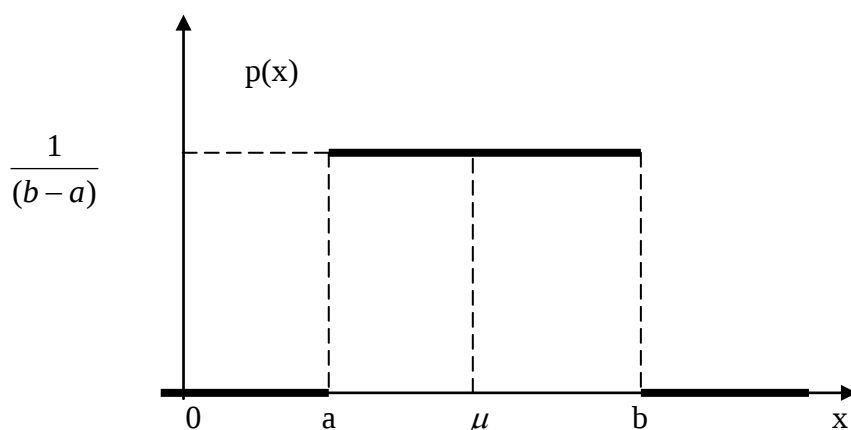
В качестве таких теоретических кривых (функций) используют хорошо изученные функции: экспоненту, логарифм, степенную функцию и их комбинации.

Простейшая функция – это константа.

5.6 Равномерное распределение

Плотность вероятности равномерного равновероятного распределения постоянна внутри интервала всех возможных значений случайной величины и равна нулю вне этого интервала.

$$p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{вне } (a, b) \end{cases}$$



Равномерное распределение часто называют прямоугольным или П-образным (из-за его характерной формы).

Среднее значение:

$$\mu = \frac{a+b}{2}$$

Дисперсия:

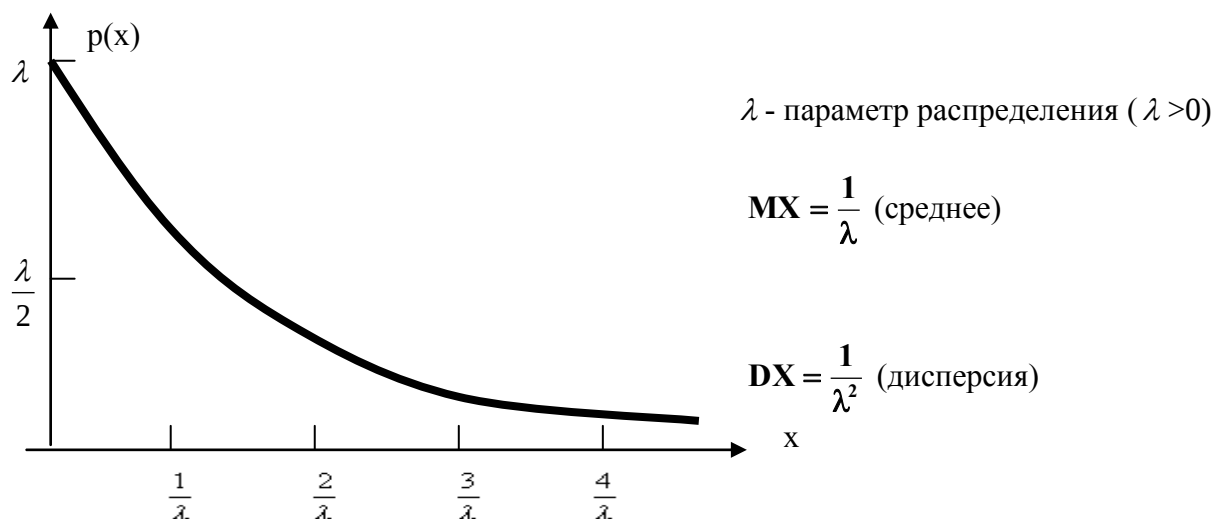
$$DX = \frac{(b-a)^2}{12}$$

С равномерным законом распределения вероятностей мы встречаемся всякий раз, когда считываем показания со шкалы измерительного прибора (например, вольтметра, микрометра, штангенциркуля и т.д.) или просто при записи чисел с конечным числом значащих цифр (т.е. при округлении чисел). При этом случайные ошибки имеют как раз равномерное распределение вероятностей.

Еще одна хорошо изученная (и любимая статистиками) функция – экспонента: e^x .

5.7 Экспоненциальное (показательное) распределение

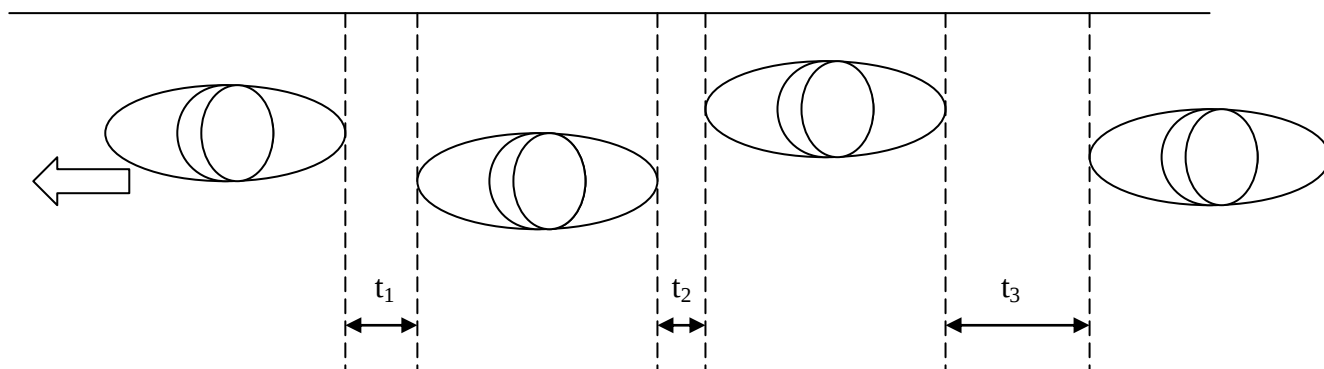
$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



Главная особенность показательного закона видна из графика его плотности распределения: вероятность появления больших и очень больших (относительно средних) значений x очень мала.

С показательным законом мы все встречаемся практически каждый день. Оказывается, что на улице с оживленным потоком автомашин интервалы времени между двумя автомобилями, движущимся друг за другом, распределены по экспоненциальному закону.

$$P(t_3) \ll P(t_1), P(t_2)$$



Это означает, что если в среднем каждая следующая автомашина появляется через $t_{cp}=30$ секунд, то вероятность того, что в потоке машин появится «разрыв» в $t=30$ минут, как правило, невелика. Это маловероятное событие, разве что где-то возник затор или произошла большая авария.

Здесь λ - это просто среднее число автомобилей, проезжающих мимо перехода в единицу времени (плотность потока автомобилей).

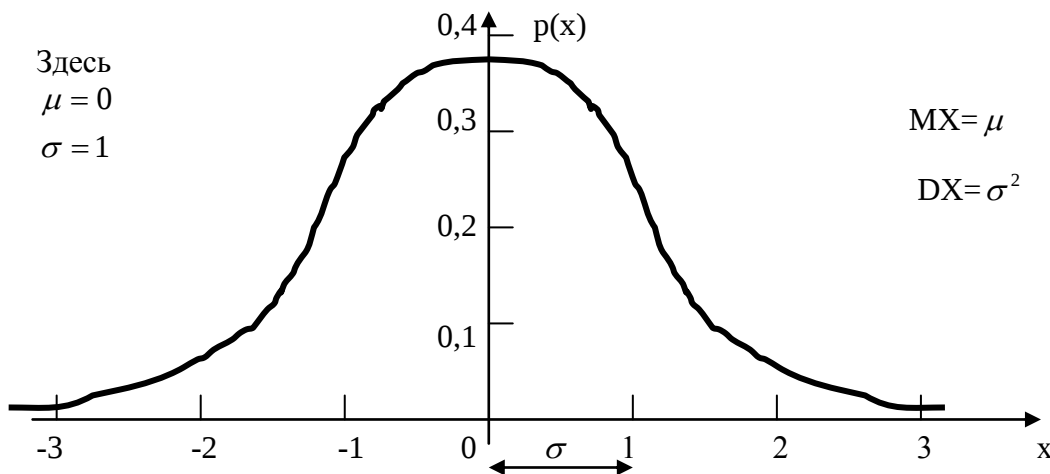
Показательное распределение вероятностей часто применяется на практике. Оно хорошо описывает, например, распад радиоактивных веществ, число вызовов абонентов на телефонной станции, время обслуживания в банке, в кассе Аэрофлота или в библиотеке и много других различных явлений.

И, наконец, самое распространенное в природе распределение вероятностей.

5.8 Распределение Гаусса или нормальное распределение

Распределение названо в честь Карла Фридриха Гаусса (1777-1855), разработавшего теорию ошибок измерений, в которой нормальное распределение играет главную роль.

$$p(x) = \frac{1}{\sigma^2 \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Гауссовское распределение занимает особое место среди всех теоретических законов распределения вероятностей. Оно описывает множество случайных процессов, и, прежде всего, ошибки любых измерений, технологических погрешностей при изготовлении деталей машин и механизмов, распределение роста и веса людей и животных и др.

Почему нормальное распределение так широко используется в практике обработки результатов измерений и оценки технологических разбросов при изготовлении деталей? Дело в том, что процесс изготовления любых деталей, например, гаек, происходит в несколько этапов и под воздействием многих независимых случайных факторов: (температуры, неоднородности инструмента, колебания напряжения в сети и т.д.)

Законы распределения этих факторов могут быть самыми различными.

Однако если вклад каждого из рассмотренных факторов не превышает вклада других, результирующее распределение технологических погрешностей оказывается нормальным! Это следствие центральной предельной теоремы теории вероятностей (Чебышев).

5.9 Свойства нормального распределения

Оно симметрично относительно μ - среднего значения, в котором находится максимум.

При $x \rightarrow \pm\infty$ плотность распределения вероятности стремится к нулю. Это значит, что вероятность больших и очень больших значений x очень мала.

Между значениями x , лежащими в интервале $(\mu - \sigma, \mu + \sigma)$ заключено 68,3% всей площади под кривой Гауссовского распределения. Это значит, что 68% всех возможных значений нормально распределенных случайных величин находится в интервале $\mu \pm \sigma$. Поэтому в технической литературе и принято указывать, что x изм. $= \mu \pm \sigma$.

99,73% всех возможных значений находится в пределах $(\mu - 3\sigma, \mu + 3\sigma)$. За пределами $\pm 3\sigma$ находится не более 0,27% возможных значений x . Поэтому на практике, если измеренное значение превышает $\pm 3\sigma$, его обычно отбрасывают как ошибочное («правило 3 сигм»).

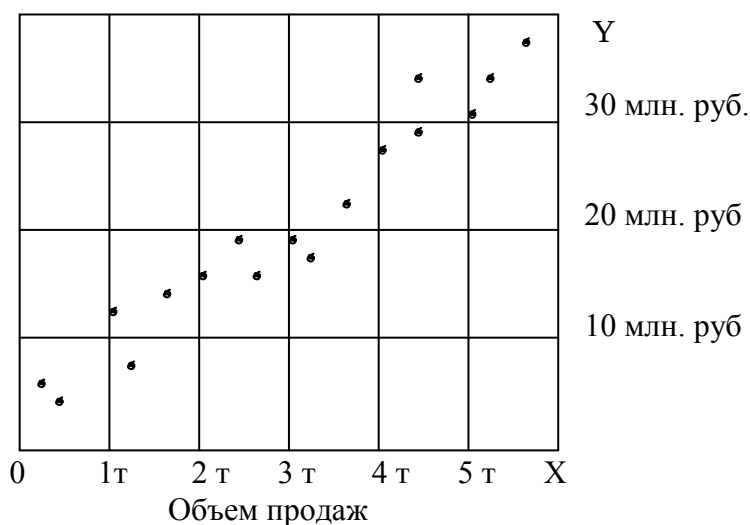
Перейдем теперь от общих вопросов теории вероятностных математических моделей к более конкретным их приложениям.

5.10 Линейная регрессия

При проведении различных измерений, опытов, экспериментов и обработке их результатов часто получают значения двух переменных y и x в виде таблицы. Например, в орошаемом земледелии y – это урожайность в центнерах с гектара, а x – количество воды, в литрах, на единицу площади. В животноводстве y – это удойность в литрах, а x – количество кормов на голову (в кг). В промышленности y может быть объемом производства, а x – прибылью от реализации продукции.

Таблица значений обычно не очень информативна. Поэтому, желая установить связь между y и x , величины y и x из таблицы наносят на график в декартовой системе координат. Как правило, точки на графике не ложатся на какую-либо хорошо известную кривую, а занимают на графике некоторую область, называемую областью рассеяния.

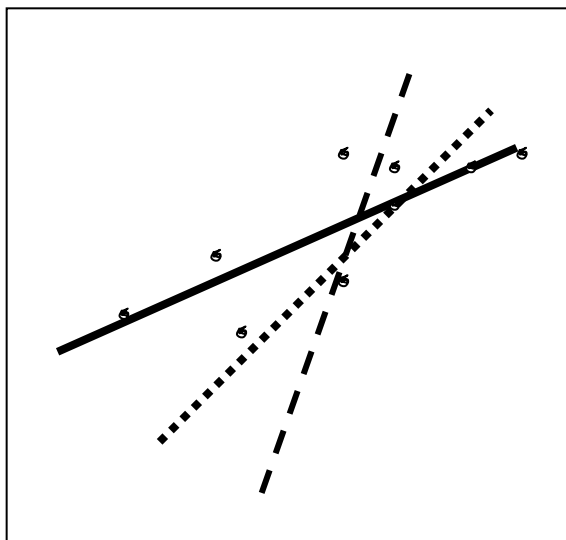
График зависимости прибыли от реализации продукции



Причины, вызывающие рассеяние, могут быть самыми разными. Это и погрешности измерений, и влияние неучтенных факторов, и различные условия, например, различные цены на сырье, и различные транспортные расходы.

Итак, на первом этапе обработке данных экспериментальные значения y и x наносят на график.

На втором этапе стараются найти плавную кривую, определяющую функциональную зависимость между y и x . Часто в качестве такой искомой функциональной зависимости можно выбрать прямую линию.



Эту прямую необходимо провести таким образом, чтобы экспериментальные точки, по возможности, располагались на одинаковом расстоянии от прямой. Но «на глаз» сделать это бывает трудно, поскольку при большом рассеянии точек прямую на графике можно сдвигать и поворачивать в широких пределах без заметного изменения ее расположения относительно экспериментальных точек.

К счастью, существует строгий математический метод, который позволяет так провести прямую, что она будет наилучшим образом отображать все экспериментальные данные.

Этот анализ называется регрессионным анализом.

Линией регрессии называется прямая линия $y = ax + b$, которая наилучшим образом аппроксимирует все экспериментальные точки в смысле метода наименьших квадратов.

Термин «регрессия» впервые ввел английский математик Фрэнсис Гальтон в 1869 г. Под регрессией он понимал движение назад по генеалогическому дереву человека («прогресс» - движение вперед, «регресс» - движение назад). Ф. Гальтон считал, что рост человека наполовину определяется ростом отца, плюс на одну четверть ростом деда, на одну восьмую ростом прадеда и т.д.

В современном понимании регрессия – это статистическая связь между случайными величинами.

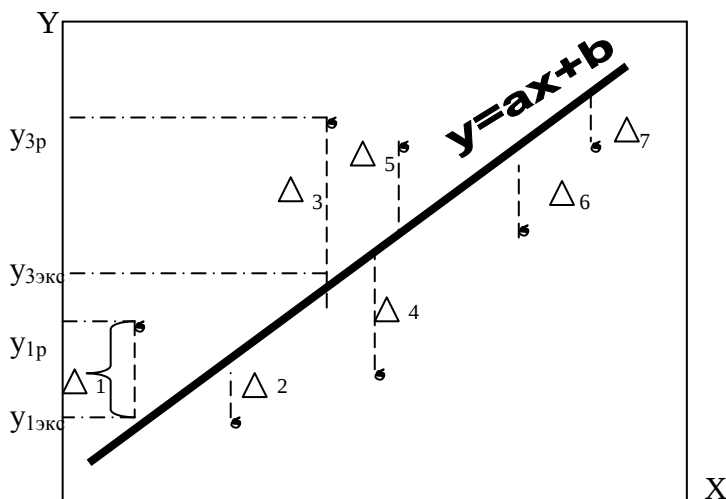
Более точно:

Регрессия – это зависимость среднего значения случайной величины от другой величины или нескольких величин.

Задача построения линии регрессии сводится к отысканию коэффициентов a и b таких, чтобы линия регрессии $y = ax + b$ проходила с наименьшим средним квадратическим отклонением от каждой из экспериментальных точек.

Эта задача решается методом наименьших квадратов. Его сущность заключается в том, что путем варьирования коэффициентов a и b подбираются такие их значения, при которых сумма квадратов отклонений расчетных значений $y_{\text{расч.}}$ от соответствующих экспериментальных значений $y_{\text{эксп.}}$ стремится к минимуму:

$$\text{Sum} = \sum_{i=1}^N (y_{\text{расч.}} - y_{\text{эксп.}})_i^2 = \sum_{i=1}^N \Delta_i^2 \rightarrow \min$$



При этом получается:

$$a = \frac{\sum_{i=1}^N x_i y_i - \frac{(\sum_{i=1}^N x_i) \cdot (\sum_{i=1}^N y_i)}{N}}{\sum_{i=1}^N x_i^2 - \frac{(\sum_{i=1}^N x_i)^2}{N}}$$

$$b = \frac{\sum_{i=1}^N y_i}{N} - a \cdot \frac{\sum_{i=1}^N x_i}{N}$$

Уравнение линейной регрессии позволяет, не проводя дополнительных измерений или экспериментов, рассчитать для заданного x величину y , т.е. прогнозировать величину y , что очень важно в тех случаях, например, при оценке предполагаемой прибыли при расширении производства.

Для количественной оценки того, насколько хорошо прямая регрессии согласуется с экспериментальными точками, вводится понятие **КОРРЕЛЯЦИИ**.

Корреляция (дословно: взаимное отношение, взаимозависимость) – это мера статистической связи случайных величин.

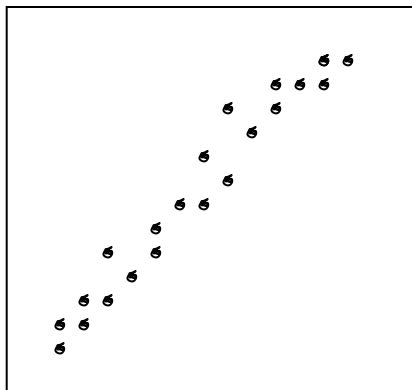
Данное понятие дает возможность судить о том, насколько тесно экспериментальные точки ложатся на прямую линейной регрессии.

Количественной мерой согласия вычисленных и измеренных значений y является коэффициент парной корреляции.

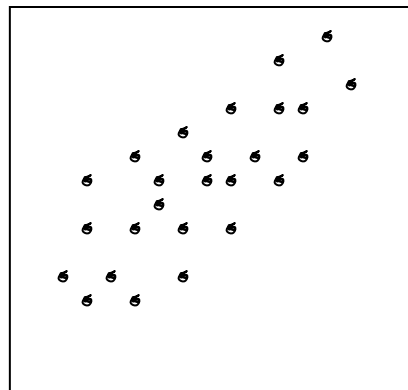
$$R^2 = \frac{\left[\sum_{i=1}^N x_i y_i - \frac{\left(\sum_{i=1}^N x_i \right) \left(\sum_{i=1}^N y_i \right)}{N} \right]^2}{\left[\sum_{i=1}^N x_i^2 - \frac{\left(\sum_{i=1}^N x_i \right)^2}{N} \right] \left[\sum_{i=1}^N y_i^2 - \frac{\left(\sum_{i=1}^N y_i \right)^2}{N} \right]}$$

Значения коэффициента парной корреляции лежат в пределах $(-1, +1)$.

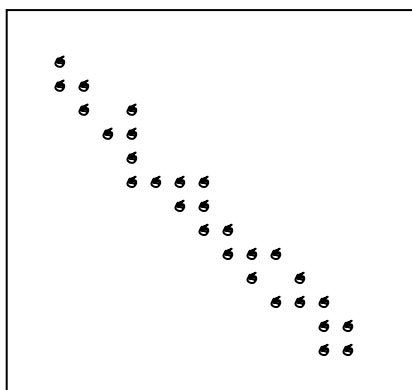
y



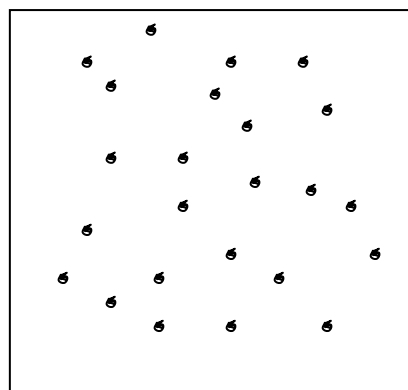
Сильная положительная
корреляция $R \approx +1$



Слабая положительная
корреляция $R \approx 0.1$



Сильная отрицательная
корреляция $R \approx -1$



Отсутствие корреляции

```
Program LinearRegression; {Построение уравнения линейной регрессии}
Const N=8;
Var a, b, S_x, S_y, S_x2, S_y2, S_xy, R:real;
    I: integer;
```

```

X: array[1..N] of real;
Y: array[1..N] of real;
Begin
  Writeln;
  Writeln (Построение уравнения линейной регрессии);
  Writeln;
  S_x:=0; S_y:=0; S_x2:=0;
  S_y2:=0; S_xy:=0;
  FOR i:=1 to n do
    Begin
      WRITE('Введите через пробел', i, '-ую точку в порядке X Y');
      READLN (X[i], Y[i]);
      S_x:=S_x+X[i];
      S_y:=S_y+Y[i];
      S_x2:=S_x2+SQR(X[i]);
      S_y2:=S_y2+SQR(Y[i]);
      S_xy:=S_xy+X[i]*Y[i];
    End;
  A:=(S_xy-S_x*S_y/n)/(S_x2-SQR(S_x)/n);
  B:=S_y/n-a*S_x/n;
  R:=SQR(S_xy-S_x*S_y/n)/((S_x2-SQR(S_x)/n)*(S_y2-SQR(S_y)/n));
  Writeln;
  Writeln('уравнение регрессии      y=';a:5:2,' x=      ',b:5:2);
  Writeln('коэффициент парной корреляции R=';R:9:7);
  Writeln;
END.

```

5.11 Прогнозирование на основе методов линейной регрессии

Одним из наиболее распространенных методов обработки экспериментальных данных при изучении различных зависимостей во многих областях науки и техники является метод линейной регрессии. Например, в сельском хозяйстве, животноводстве, в производстве и распределении часто возникают задачи нахождения зависимостей, наилучшим образом выражающих связь между урожайностью и количеством поливов в орошаемом земледелии, величиной удоя и количеством кормов, прибылью и объемам производства или реализации товара и т.д.

В рамках регрессионного анализа для наилучшего подбора аппроксимирующей кривой к экспериментальным точкам используется метод наименьших квадратов, который позволяет свести к минимуму отклонение экспериментальных точек от подбираемой кривой. Для оценки того, насколько хорошо кривая согласуется с экспериментальными данными, вводится понятие коэффициента корреляции. Это понятие дает возможность судить о том, насколько тесно ложатся экспериментальные точки на аппроксимирующую кривую.

При проведении различных измерений, экспериментов и обработке их результатов часто получают значения двух переменных в виде таблицы или набора точек на плоскости в декартовой системе координат. Обычно точки не ложатся на какую-либо прямую, а занимают на графике некоторую область, называемую областью рассеяния.

Для определения функциональной зависимости y от x через эти точки необходимо бывает провести прямую линию и получить ее уравнение в форме:

$$y = ax + b$$

Прямую необходимо провести так, чтобы измеренные точки по возможности располагались на одинаковом расстоянии от прямой. На глаз это бывает сделать трудно, т.к. прямая

может сдвигаться и поворачиваться на плоскости в некоторых пределах без заметного на глаз изменения ее расположения относительно экспериментальных точек.

Метод линейной регрессии позволяет построить прямую так, что она будет проходить с наименьшим среднеквадратическим отклонением от каждой из измеренных точек. Задача определения коэффициентов a и b , входящих в уравнение регрессии, решается методом наименьших квадратов.

Уравнение регрессии позволяет, не проводя дальнейших экспериментов, рассчитать для данного x величину y , т.е. прогнозировать величину y . Такой прогноз, разумеется, не будет абсолютно точным; близость прогнозируемой величины y к фактической будет тем лучше, чем точнее зависимость y от x ближе к линейной. Количественной мерой согласия вычисленных и измеренных значений y является коэффициент корреляции.

Типовой пример. Промышленное предприятие производит в месяц X тонн продукции. Доход от реализации этой продукции Y зависит от многих различных факторов, например, объема продаж, транспортных расходов, стоимости сырья и так далее.

В месяц предприятие производит от 1 до 8 тонн продукции. Данные по объему производства и прибыли представлены в Таблице.

№	1	2	3	4	5	6	7	8
X	1	2	3	4	5	6	7	8
Y	148.8	160.4	167.6	170.2	176.8	197.3	198.0	198.7

После модернизации оборудования предполагается увеличить объем производства продукции в 1.5 раза. Требуется на основе данных по объему производства и прибыли, представленных в Таблице, сделать прогноз по увеличению прибыли предприятия.

Прямая линейной регрессии, полученная по экспериментальным данным, имеет вид:

$$Y = 7,54 X + 143,31$$

Коэффициент парной корреляции:

$$R^2 = 0,943076$$

Подставляя в полученную формулу для y значение $X=12$, получаем $Y=233.79$, т.е. при увеличении объема производства в 1.5 раза прибыль предприятия должна возрасти примерно в 1.3 раза.

Коэффициент корреляции близок к 1, что свидетельствует о хорошем согласии экспериментальных и расчетных данных.

Вопросы для самоконтроля

1. Какие математические модели называют вероятностными.
2. В чём состоит различие между дискретными и непрерывными случайными величинами
3. Что принято называть распределением вероятности случайной величины
4. Каковы свойства вероятностей случайных величин?
5. Назовите наиболее часто используемые числовые характеристики случайных величин
6. Что называют законом распределения вероятностей непрерывной случайной величины
7. Как определяется среднее значение и дисперсия случайной величины
8. Что называют модельными законами распределения вероятности
9. Основные свойства нормального распределения

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд., стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 6

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ

6.1. Общие положения

Проблемы, с которыми приходится сталкиваться в различных сферах нашей жизни (в промышленном производстве, сельском хозяйстве, торговле, управлении и т. д.), постоянно усложняются. Это определяет необходимость совершенствования имеющихся и разработки новых методов и процедур их решения. В этой лекции показано, что имитационное моделирование является эффективным средством решения сложных проблем. Имитационные модели могут применяться для:

- исследования границ и структур систем с целью решения конкретных проблем;
- определения и анализа критических, элементов, компонентов и точек в исследуемых системах и процессах;
- синтеза и оценки предполагаемых решений;
- прогнозирования и планирования будущего развития исследуемых систем.

Процесс имитационного моделирования начинается с определения подлежащих решению проблем, что в свою очередь определяет состав и границы исследуемой системы. Построение имитационной модели исследуемой системы, хотя и зависит от специфики решаемой проблемы, требует определенной методологической схемы. Имитационный язык обеспечивает исследователя такой схемой, а также осуществляет трансляцию модели в доступную вычислительной системе форму. Компьютер, на котором исследуется разработанная имитационная модель, выдает информацию о поведении модели, которая затем может анализироваться в процессе решения проблемы.

Цель лекции — ознакомить обучающихся с основами имитационного моделирования, и практикой его использования для решения конкретных проблем. В данной лекции представлены основные понятия имитационного моделирования и этапы имитационного исследования.

6.2. Виды моделей

Модель — это описание системы. В физических науках модели обычно разрабатываются на основе теоретических законов и принципов. Моделями могут быть масштабированные физические объекты (иконические модели), математические уравнения и взаимосвязи (абстрактные модели) или графические модели (визуальные модели). Модели эффективно используются, в частности, при описании, проектировании и анализе систем.

Разработка модели — сложный процесс, который во многом является искусством, но, однако, упрощается, если:

- 1) известны физические законы, описывающие функционирование системы;
- 2) может быть разработано графическое представление системы;
- 3) можно управлять входами, элементами и выходами системы [7].

Моделирование сложных крупномасштабных систем чаще всего представляет собой более трудную задачу, чем моделирование физических систем. Это объясняется следующими причинами:

- 1) в распоряжении исследователя имеется мало фундаментальных законов, относящихся к рассматриваемой системе;
- 2) многие взаимосвязи между элементами в системе с трудом поддаются количественному описанию и формализации;
- 3) трудно количественно описать поведение входных элементов;

- 4) важную роль играют стохастические процессы;
- 5) неотъемлемой частью таких систем является процесс принятия решений человеком.

6.3. Построение моделей

Модель представляет собой абстрактное описание системы, уровень детализации которого определяет сам исследователь! Человек принимает решение о том, является ли данный элемент системы существенным, а, следовательно, будет ли он включен в описание системы. Это решение принимается с учетом цели, лежащей в основе разработки модели. От того, насколько хорошо исследователь умеет выделять существенные элементы и взаимосвязи между ними, зависит успех моделирования.

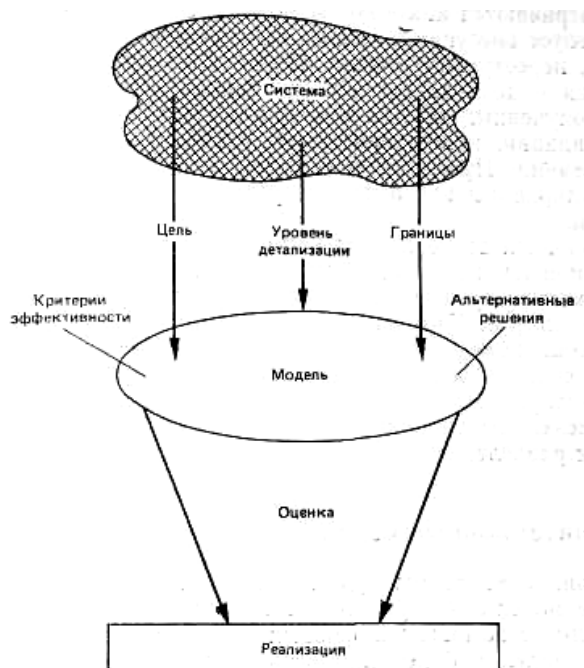


Рис. 6.1. Процесс построения модели.

На рис. 6.1 приведена схема предлагаемого подхода к построению моделей. Система рассматривается как состоящая из множества взаимосвязанных элементов, объединенных для выполнения определенной функции. Определение системы во многом субъективно, т. е. оно зависит не только от цели разработки модели, но и от того, кто именно определяет систему.

Итак, процесс моделирования начинается с определения цели разработки модели, на основе которой затем устанавливаются границы системы и необходимый уровень детализации моделируемых процессов. Выбранный уровень детализации должен позволять абстрагироваться от неточно определённых из-за недостатка информации аспектов функционирования реальной системы. В описание системы, кроме того, должны быть включены критерии эффективности функционирования системы и оцениваемые альтернативные решения, которые могут рассматриваться как часть модели или как ее входы. Оценки же альтернативных решений по заданным критериям эффективности рассматриваются как выходы модели. Обычно оценка альтернатив требует внесения изменений в описание системы и, следовательно, перестройки модели. Поэтому на практике процесс построения модели является итеративным. После того как на основе полученных оценок альтернатив могут быть выработаны рекомендации, можно приступить к вне-

дрению результатов моделирования. При этом в рекомендациях должны быть четко сформулированы как основные решения, так и условия их реализации.

Описанный выше подход полностью применим к построению имитационных моделей. С помощью имитационного моделирования могут строиться как агрегированные, так и детализированные модели. Имитационному моделированию также свойственна концепция итеративного построения модели, в ходе которого модель изменяется путем добавления новых или исключения некоторых ее элементов и (или) взаимосвязей между ними. Эти аспекты имитационного моделирования описываются в следующем разделе.

6.4. Имитационные модели

В данной лекции мы будем рассматривать только те имитационные модели, которые разрабатываются для проведения исследований с помощью цифровых ЭВМ. Таким образом, имитационной моделью будем называть логико-математическое описание системы, которое может быть исследовано в ходе проведения экспериментов на цифровой ЭВМ и, следовательно, может считаться лабораторной версией системы. После окончания разработки имитационной модели с ней проводятся машинные эксперименты, которые позволяют сделать выводы о поведении системы:

- без ее построения, если это проектируемая система;
- без вмешательства в ее функционирование, если это действующая система, экспериментирование с которой или слишком дорого, или небезопасно;
- без ее разрушения, если цель эксперимента состоит в определении пределов воздействия на систему.

Таким образом, имитационные модели могут использоваться для проектирования, анализа и оценки функционирования систем. В настоящее время имитационное моделирование используется для исследования разнообразных систем, в частности городских, экономических, коммерческих, производственных, биологических, социальных, транспортных систем, систем здравоохранения и др. Наряду со статистическими, имитационные методы являются наиболее распространенными средствами теории управления и исследования операций в управлении промышленными и торговыми предприятиями и организациями.

В имитационном моделировании предполагается, что систему можно описать в терминах, понятных вычислительной системе. Ключевым моментом при этом является выделение и описание состояний системы. Система характеризуется набором переменных, каждая комбинация значений которых описывает ее конкретное состояние. Следовательно, путем изменения значений переменных можно имитировать переход системы из одного состояния в другое. Таким образом, имитационное моделирование — это представление динамического поведения системы посредством продвижения ее от одного состояния к другому в соответствии с хорошо определенными операционными правилами.

Изменения состояния системы могут происходить либо непрерывно, либо в дискретные моменты времени. Хотя процедуры описания динамического поведения дискретно и непрерывно изменяющихся моделей различны, основная концепция имитации системы — отображение изменений ее состояния с течением времени — остается той же. В следующем разделе мы проиллюстрируем это положение на примере простой дискретно изменяющейся системы.

6.5. Пример имитационного моделирования

В качестве примера имитационного моделирования рассмотрим обслуживание кассиром клиентов в банке. Клиенты прибывают в банк, ожидают обслуживания, если кассир занят, обслуживаются и затем покидают банк. Клиенты, прибывающие в банк в тот

момент, когда кассир занят, ожидают в единственной к нему очереди. Для простоты мы предполагаем, что момент прибытия клиента и время обслуживания его кассиром известны для каждого клиента (табл. 6.1). Наша цель состоит в имитации функционирования описанной выше системы для того, чтобы определить, какой процент времени кассир свободен и каково среднее время пребывания клиента в банке.

Так как имитационное моделирование — это динамическое отображение изменений состояния системы с течением времени, то в первую очередь должны быть определены состояния системы. Для данного примера состояние моделируемой системы может определяться состоянием кассира (свободен или занят) и числом клиентов в банке.

Таблица 6.1

Моменты прибытия и времена обслуживания клиентов

Номер клиента	Момент прибытия, мин	Время обслуживания, мин
1	3,2	3,8
2	10,9	3,5
3	13,2	4,2
4	14,8	3,1
5	17,7	2,4
6	19,8	4,3
7	21,5	2,7
8	26,3	2,1
9	32,6	2,5
10	36,6	3,4

Состояние системы изменяется в результате:

- 1) прибытия клиента в банк;
- 2) завершения обслуживания кассиром и последующего ухода клиента.

Для иллюстрации имитации мы будем определять состояние системы путем обработки упорядоченных во времени событий, соответствующих прибытию и уходу клиента.

Итоги имитации в соответствии с исходными данными, представленными в табл. 6.1, приводятся в табл. 6.2. При этом предполагалось, что в начальный момент времени в системе нет клиентов, кассир свободен и первый клиент прибывает в момент времени, равный 3,2 мин.

В табл. 6.2 первый и второй столбцы взяты из табл. 6.1. Время начала обслуживания, приведенное в третьем столбце, зависит от того, покинул ли предыдущий клиент банк. Оно принимается равным наибольшему значению из времен прибытия данного клиента и ухода предыдущего. Время ухода, приведенное в четвертом столбце, вычисляется как сумма соответствующего элемента третьего столбца и времени обслуживания данного клиента, которое определяется по табл. 6.1. Значение времени нахождения каждого клиента в очереди и в банке вычисляется, как показано в табл. 6.2. Средние значения этих переменных соответственно равны 2,61 и 5,81 мин.

Моделирование обслуживания клиентов банка

Номер клиента (1)	Момент прибытия, мин (2)	Момент начала обслуживания, мин (3)	Момент ухода, мин (4)	Время ожидания в очереди, мин (5)=(3)-(2)	Время пребывания в банке, мин (6)=(4)-(2)
1	2	3	4	5	6
1	3,2	3,2	7,0	0,0	3,8
2	10,9	10,9	14,4	0,0	3,5
3	13,2	14,4	18,6	1,2	5,4
4	14,8	18,6	21,7	3,8	6,9
5	17,7	21,7	24,1	4,0	6,4
6	19,8	24,1	28,4	4,3	8,6
7	21,5	28,4	31,1	6,9	9,6
8	26,3	31,1	33,2	4,8	6,9
9	32,6	33,2	35,7	1,1	3,6
10	36,6	36,6	40,0	0,0	3,4

Таблица 6.2 содержит итоговую информацию, относящуюся к клиенту, но не содержит сведений о кассире и размере очереди к нему. Для получения такой информации необходимо исследовать события, связанные с этой ситуацией.

Логика обработки событий прибытия и ухода зависит от состояния системы в момент наступления этих событий. При наступлении события «прибытие клиента» в банк дальнейшая ситуация определяется состоянием кассира. Если кассир свободен, он переходит в состояние «занят» и приступает к обслуживанию клиента. При этом планируется событие «уход данного клиента в момент времени, равный текущему времени плюс продолжительность его обслуживания. Если же кассир занят, обслуживание клиента не может начаться, и, следовательно, он встает в очередь (длина очереди увеличивается на 1). Логика обработки события «уход клиента из банка» зависит от длины очереди. Если в очереди есть хотя бы один клиент, кассир остается в состоянии «занят», длина очереди уменьшается на 1 и для первого клиента из очереди планируется событие ухода. Если же очередь пуста, кассир переводится в состояние «свободен».

В табл. 6.1 приводится событийно-ориентированное описание состояния кассира и числа клиентов в банке (события расположены в хронологическом порядке). На рис. 6.2 приведены графики изменения значений этих переменных состояний во времени. Результаты имитации показывают, что в течение первых 40 мин работы в банке в среднем одновременно находилось 1,4525 клиента, а кассир был свободен 20% времени.

Для расположения событий прихода и ухода в хронологическом порядке необходимо вести запись (календарь) событий, подлежащих последующей обработке (будущих событий). Это осуществляется путем фиксирования моментов наступления следующего события прихода и следующего события ухода. Сравнение этих моментов определяет затем выбор одного из событий для обработки. Такой упорядоченный список событий обычно называется файлом или календарем событий.

Событийно-ориентированное описание работы банковского кассира

Время события, мин	Номер клиента	Тип события	Длина очереди	Число клиентов	Состояние кассира	Время простоя кассира, мин
0,0	-	Начало	0	0	Свободен	-
3,2	1	Приход	0	1	Занят	3,2
7,0	1	Уход	0	0	Свободен	
10,9	2	Приход	0	1	Занят	3,9
13,2	3	Приход	1	2	Занят	
14,4	2	Уход	0	1	Занят	

Время события, мин	Номер клиента	Тип события	Длина очереди	Число клиентов	Состояние кассира	Время простоя кассира, мин
14,8	4	Приход	1	2	Занят	
17,7	5	Приход	2	3	Занят	
18,6	3	Уход	1	2	Занят	
19,8	6	Приход	2	3	Занят	
21,5	7	Приход	3	4	Занят	
21,7	4	Уход	2	3	Занят	
24,1	5	Уход	1	2	Занят	
26,3	8	Приход	2	3	Занят	
28,4	6	Уход	1	2	Занят	
31,1	7	Уход	0	1	Занят	
32,1	9	Приход	1	2	Занят	
33,2	8	Уход	0	1	Занят	
35,7	9	Уход	0	0	Свободен	
36,6	10	Приход	0	1	Занят	0,9
40,0	10	Уход	0	0	Свободен	

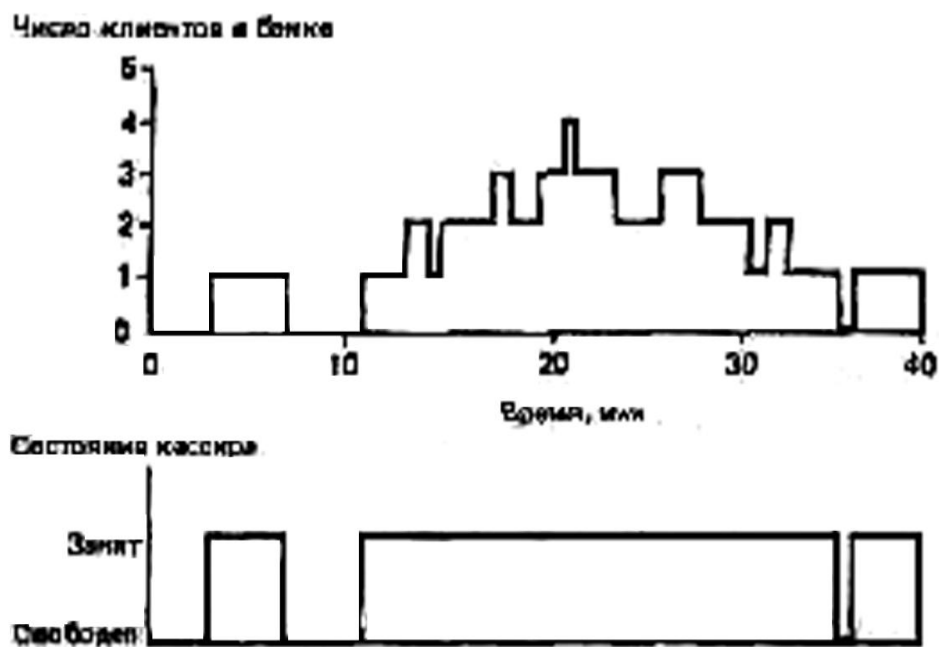


Рис. 6.2. Графическое представление результатов моделирования работы банка.

Приведенный пример иллюстрирует целый ряд важных понятий. Мы убедились в том, что в любой момент имитационного времени модель находится в некотором состоянии. При наступлении событий состояние системы меняется в соответствии с логико-математическими отношениями, связанными с этими событиями. Таким образом, события определяют динамическую структуру системы. При заданных начальном состоянии, логической схеме обработки каждого события и методе определения случайных величин наша проблема в большей степени становится вычислительной проблемой. Существенным элементом схемы вычислений является календарь событий, который обеспечивает механизм для записи и упорядочения будущих событий.

Еще одним важным моментом является то, что мы можем рассматривать изменения состояния системы в двух аспектах:

- 1) как процесс, с которым сталкивается клиент при обслуживании (точка зрения клиента),
- 2) как последовательность событий, которые вызывают изменение состояния кассира (точка зрения кассира).

6.6. Этапы имитационного моделирования

Как уже говорилось выше, процесс последовательной разработки имитационной модели начинается с создания простой модели, которая затем постепенно усложняется в соответствии с требованиями, предъявляемыми решаемой проблемой. В процессе имитационного моделирования можно выделить следующие основные этапы:

1. **Формулирование проблемы:** описание исследуемой проблемы и определение целей исследования.
2. **Разработка модели:** логико-математическое описание моделируемой системы в соответствии с формулировкой проблемы.
3. **Подготовка данных:** идентификация, спецификация и сбор данных.
4. **Трансляция модели:** перевод модели на язык, приемлемый для используемой ЭВМ.
5. **Верификация:** установление правильности машинных программ.
6. **Валидация:** оценка требуемой точности и соответствия имитационной модели реальной системе.
7. **Стратегическое и тактическое планирование:** определение условий проведения машинного эксперимента с имитационной моделью.
8. **Экспериментирование:** прогон имитационной модели на ЭВМ для получения требуемой информации.
9. **Анализ результатов:** изучение результатов имитационного эксперимента для подготовки выводов и рекомендаций по решению проблемы.
10. **Реализация и документирование:** реализация рекомендаций, полученных на основе имитации, и составление документации по модели и ее использованию.

Хотя некоторые из этих этапов уже обсуждались выше в связи с построением модели, мы, тем не менее, снова кратко рассмотрим основные этапы процесса имитационного моделирования.

Первой задачей имитационного моделирования является точное определение проблемы и детальная формулировка целей исследования. Как правило, определение проблемы является непрерывным процессом, который обычно осуществляется в течение всего исследования. Оно пересматривается по мере более глубокого понимания исследуемой проблемы и возникновения новых ее аспектов.

Как только сформулировано начальное определение проблемы, начинается этап построения модели исследуемой системы. Модель включает статистическое и динамическое описание системы. В статистическом описании определяются элементы системы и их характеристики, а в динамическом — взаимодействия элементов системы, в результате которых происходят изменения ее состояния во времени.

Процесс формулирования модели во многом является искусством. Разработчик модели должен понять структуру системы, выявить правила ее функционирования и суметь выделить в них самое существенное, исключив ненужные детали. Модель должна быть простой для понимания и в то же время достаточно сложной, чтобы реалистично отображать характерные черты реальной системы. Наиболее важными являются принимаемые разработчиком решения относительно того, верны ли принятые упрощения и допущения, какие элементы и взаимодействия между ними должны быть включены в модель. Уровень детализации модели зависит от цели ее создания. Необходимо рассматривать только те элементы, которые имеют существенное значение для решения исследуемой проблемы. Как на этапе формулирования проблемы, так и на этапе моделирования необходимо тесное взаимодействие между разработчиками модели и ее пользователями. «Первый эскиз» модели должен быть построен, проанализирован и обсужден. Во многих случаях это требует от разработчиков большой ответственности и готовности продемонстрировать свое возможное незнание исследуемой системы. Однако эволюционный процесс моделирования позволяет быстрее обнаруживать допущен-

ные разработчиками неточности и более эффективно их конкретизировать. Кроме того, тесное взаимодействие на этапах формулирования проблемы и разработки модели создает у пользователя уверенность в правильности модели и поэтому помогает обеспечить успешную реализацию результатов имитационного исследования.

На этапе разработки модели определяются требования к входным данным. Некоторые из этих данных могут уже быть в распоряжении разработчика модели, в то время как для сбора других потребуются время и усилия. Обычно значения таких входных данных задаются на основе некоторых гипотез или предварительного анализа. В некоторых случаях точные значения одного (и более) входных параметров оказывают небольшое влияние на результаты прогонов модели. Чувствительность получаемых результатов к изменению входных данных может быть оценена путем проведения серии имитационных прогонов для различных значений входных параметров. Имитационная модель, следовательно, может использоваться для уменьшения затрат времени и средств на уточнение входных данных.

После того как разработана модель и собраны начальные входные данные, следующей задачей является перевод модели в форму, доступную для ЭВМ. Хотя для программирования имитационной модели может использоваться универсальный язык, применение специализированного имитационного языка имеет существенные преимущества. Помимо сокращения времени программирования использование имитационного языка упрощает разработку модели, так как язык содержит набор понятий для формализованного описания системы. Для указанных целей обычно применяют язык СЛАМ [SLAM (Simulation Language for Alternative Modeling) — имитационный язык для альтернативного моделирования], который оснащен графическими средствами для объединения в единый процесс разработки модели и ее трансляции. СЛАМ также содержит программные средства для последовательной модификации моделей на любом уровне детализации, требуемой для отображения сложности исследуемой системы.

На этапах верификации и валидации осуществляется оценка функционирования имитационной модели. На этапе верификации определяется, соответствует ли запрограммированная для ЭВМ модель замыслу разработчика. Это обычно осуществляется путем ручной (аналитической) проверки вычислений, а также может быть использован и ряд статистических методов.

Установление адекватности имитационной модели исследуемой системе осуществляется на этапе валидации. Валидация модели обычно выполняется на различных уровнях. Мы рекомендуем выполнять валидацию на уровне входных данных, элементов модели, подсистем и их взаимосвязей. И хотя валидация имитационных моделей достаточно сложна, она является существенно более легкой задачей, чем валидация моделей других типов, например моделей линейного программирования. В имитационных моделях существует соответствие между элементами модели и элементами реальной системы, поэтому проверка адекватности разработанной модели включает сравнение ее структуры со структурой системы, а также сравнение того, как реализованы элементарные функции и решения в модели и системе.

Специальные методы валидации включают установление адекватности путем использования постоянных значений всех параметров имитационной модели или путем оценивания чувствительности выходов к изменению значений входных данных. В процессе валидации сравнение должно осуществляться на основе анализа как реальных, так и экспериментальных данных о функционировании системы. Следует помнить, что имеющиеся в распоряжении исследователя реальные данные о функционировании системы являются всего лишь выборкой из того, что могло произойти в прошлом.

Условия проведения машинных прогонов модели определяются на этапах стратегического и тактического планирования. Задача стратегического планирования заключается в разработке эффективного плана эксперимента, в результате которого либо выясняется взаимосвязь между управляемыми переменными, либо находится комбинация

значений управляемых переменных, минимизирующая или максимизирующая отклик имитационной модели. В тактическом планировании в отличие от стратегического решается вопрос о том, как в рамках плана эксперимента провести каждый имитационный прогон, чтобы получать наибольшее количество информации из выходных данных. Важное место в тактическом планировании занимают определение начальных условий имитационных прогонов и методы снижения дисперсии среднего значения отклика модели.

Следующие этапы в процессе имитационного исследования — проведение машинного эксперимента и анализ результатов — включают прогон имитационной модели на компьютере и интерпретацию полученных выходных данных. При использовании результатов имитационных экспериментов для подготовки выводов или проверки гипотез о функционировании реальной системы применяются статистические методы.

Последним этапом в процессе имитационного исследования являются реализация полученных решений и документирование имитационной модели и ее использования. Ни один из имитационных проектов не должен считаться законченным до тех пор, пока их результаты не были использованы в процессе принятия решений. Успех реализации во многом зависит от того, насколько правильно разработчик модели выполнил все предыдущие этапы процессов имитационного исследования. Если разработчик и пользователь работали в тесном контакте и достигли взаимопонимания при разработке модели и ее исследовании, то результаты проекта, скорее всего, будут успешно внедряться. Если же между ними не было тесной взаимосвязи, то, несмотря на элегантность и адекватность имитационной модели, сложно будет разработать эффективные рекомендации.

Названные выше этапы имитационного исследования редко выполняются в строго заданной последовательности, начиная с определения проблемы и кончая документированием. В ходе имитационного исследования могут быть сбои в прогонах модели, ошибочные допущения, от которых в дальнейшем приходится отказываться, переформулировки целей исследования, повторные оценки и перестройки модели. Такой итеративный процесс позволяет разработать имитационную модель, которая дает верную оценку альтернатив и облегчает процесс принятия решения.

6.7. Выводы

Имитационное моделирование — это эффективное средство, которое широко используется для решения сложных проблем, а имитационная модель — абстрактное описание системы. Логика построения модели должна быть простой и понятной всем участникам имитационного исследования; в случае необходимости имитационная модель должна легко модифицироваться. Реализация полученных с помощью имитации рекомендаций является неотъемлемой частью имитационной методологии.

Вопросы для самоконтроля

1. Что принято называть имитационной математической моделью
2. Для каких целей применяются имитационные математические модели.
3. В чём состоит процесс имитационного моделирования
4. Принципы построения имитационных математических моделей
5. Основные этапы имитационного моделирования
6. Валидация и верификация имитационной математической модели
7. В каких случаях применение имитационного моделирования наиболее эффективно
8. Как определяются состояния моделируемой системы в процессе её имитации

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Data mining. КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ

7.1. Современные технологии поиска и анализа данных

Выделяют пять основных методов поиска и анализа данных и полезной информации: ассоциация, последовательность, классификация, кластеризация и прогнозирование.

Ассоциация имеет место в том случае, если несколько событий связаны друг с другом. Например, исследование, проведенное в супермаркете, может показать, что 65% купивших кукурузные чипсы берут также и "кока-колу", а при наличии скидки за такой комплект "колу" приобретают в 85% случаев. Располагая сведениями о подобной ассоциации, менеджерам легко оценить, насколько действенна предоставляемая скидка.

Если существует цепочка связанных во времени событий, то говорят о последовательности. Так, например, после покупки дома в 45% случаев в течение месяца приобретается и новая кухонная плита, а в пределах двух недель 60% новоселов обзаводятся холодильником.

С помощью классификации выявляются признаки, характеризующие группу, к которой принадлежит тот или иной объект. Это делается посредством анализа уже классифицированных объектов и формулирования некоторого набора правил.

Кластеризация отличается от классификации тем, что сами группы заранее не заданы. С помощью кластеризации средства Data Mining самостоятельно выделяют различные однородные группы данных.

Основой для всевозможных систем прогнозирования служит историческая информация, хранящаяся в БД в виде временных рядов. Если удастся построить найденные шаблоны, адекватно отражающие динамику поведения целевых показателей, есть вероятность, что с их помощью можно предсказать и поведение системы в будущем.

7.2. Технология Data Mining: понятие, назначение

Data Mining переводится как "добыча" или "раскопка данных". Нередко рядом с Data Mining встречаются слова "обнаружение знаний в базах данных" (knowledge discovery in databases) и "интеллектуальный анализ данных". Их можно считать синонимами Data Mining. Возникновение всех указанных терминов связано с новым витком в развитии средств и методов обработки данных.

До начала 90-х годов, казалось, не было особой нужды переосмысливать ситуацию в этой области. Все шло своим чередом в рамках направления, называемого прикладной статистикой. Теоретики проводили конференции и семинары, писали внушительные статьи и монографии, изобиловавшие аналитическими выкладками.

В связи с совершенствованием технологий записи и хранения данных на людей обрушились колоссальные потоки информационной руды в самых различных областях. Деятельность любого предприятия (коммерческого, производственного, медицинского, научного и т.д.) теперь сопровождается регистрацией и записью всех подробностей его деятельности. Что делать с этой информацией? Стало ясно, что без продуктивной переработки потоки сырых данных образуют никому не нужную свалку. Специфика современных требований к такой переработке следующие:

- Данные имеют неограниченный объем
- Данные являются разнородными (количественными, качественными, текстовыми)
- Результаты должны быть конкретны и понятны

- Инструменты для обработки сырых данных должны быть просты в использовании

Традиционная математическая статистика, долгое время претендовавшая на роль основного инструмента анализа данных, перед лицом возникших проблем оказалась малоэффективной. Главная причина — концепция усреднения по выборке, приводящая к операциям над фиктивными величинами (типа средней температуры пациентов по больнице, средней высоты дома на улице, состоящей из дворцов и лачуг и т.п.). Методы математической статистики оказались полезными главным образом для проверки заранее сформулированных гипотез (verification-driven data mining) и для “грубого” разведочного анализа, составляющего основу оперативной аналитической обработки данных (online analytical processing, OLAP).

В основу современной технологии Data Mining (discovery-driven data mining) положена концепция шаблонов (паттернов), отражающих фрагменты многоаспектных взаимоотношений в данных. Эти шаблоны представляют собой закономерности, свойственные подвыборкам данных, которые могут быть компактно выражены в понятной человеку форме. Поиск шаблонов производится методами, не ограниченными рамками априорных предположений о структуре выборке и виде распределений значений анализируемых показателей. Важное положение Data Mining — нетривиальность разыскиваемых шаблонов. Это означает, что найденные шаблоны должны отражать неочевидные, неожиданные (unexpected) регулярности в данных, составляющие так называемые скрытые знания (hidden knowledge). К обществу пришло понимание, что сырые данные (raw data) содержат глубинный пласт знаний, при грамотной раскопке которого могут быть обнаружены настоящие самородки.

В целом технологию Data Mining достаточно точно определяет Григорий Пиатецкий-Шапиро — один из основателей этого направления:

Data Mining — это процесс обнаружения в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

Сфера применения Data Mining ничем не ограничена — она везде, где имеются какие-либо данные. Но в первую очередь методы Data Mining сегодня заинтересовали коммерческие предприятия, развертывающие проекты на основе информационных хранилищ данных (Data Warehousing). Опыт многих таких предприятий показывает, что отдача от использования Data Mining может достигать 1000%. Например, известны сообщения об экономическом эффекте, в 10–70 раз превысившем первоначальные затраты от 350 до 750 тыс. дол. Известны сведения о проекте в 20 млн. дол., который окупился всего за 4 месяца. Другой пример — годовая экономия 700 тыс. дол. за счет внедрения Data Mining в сети универсамов в Великобритании.

Data Mining представляют большую ценность для руководителей и аналитиков в их повседневной деятельности. Деловые люди осознали, что с помощью методов Data Mining они могут получить ощутимые преимущества в конкурентной борьбе. Кратко охарактеризуем некоторые возможные бизнес-приложения Data Mining [2].

7.3. Приложения технологии Data Mining

- Коммерция
- Банковское дело
- Телекоммуникации
- Страхование
- Другие

Коммерция. Предприятия розничной торговли сегодня собирают подробную информацию о каждой отдельной покупке, используя кредитные карточки с маркой магазина и компьютеризованные системы контроля. Вот типичные задачи, которые можно решать с помощью Data Mining в сфере розничной торговли:

- анализ покупательской корзины (анализ сходства) предназначен для выявления товаров, которые покупатели стремятся приобретать вместе. Знание покупательской корзины необходимо для улучшения рекламы, выработки стратегии создания запасов товаров и способов их раскладки в торговых залах.
- исследование временных шаблонов помогает торговым предприятиям принимать решения о создании товарных запасов. Оно дает ответы на вопросы типа "Если сегодня покупатель приобрел видеокамеру, то через какое время он вероятнее всего купит новые батарейки и пленку?"
- создание прогнозирующих моделей дает возможность торговым предприятиям узнавать характер потребностей различных категорий клиентов с определенным поведением, например, покупающих товары известных дизайнеров или посещающих распродажи. Эти знания нужны для разработки точно направленных, экономичных мероприятий по продвижению товаров.

Банковское дело. Достижения технологии Data Mining используются в банковском деле для решения следующих распространенных задач:

- выявление мошенничества с кредитными карточками. Путем анализа прошлых транзакций, которые впоследствии оказались мошенническими, банк выявляет некоторые стереотипы такого мошенничества.
- сегментация клиентов. Разбивая клиентов на различные категории, банки делают свою маркетинговую политику более целенаправленной и результативной, предлагая различные виды услуг разным группам клиентов.
- прогнозирование изменений клиентуры. Data Mining помогает банкам строить прогнозные модели ценности своих клиентов, и соответствующим образом обслуживать каждую категорию.

Телекоммуникации. В области телекоммуникаций методы Data Mining помогают компаниям более энергично продвигать свои программы маркетинга и ценообразования, чтобы удерживать существующих клиентов и привлекать новых. Среди типичных мероприятий отметим следующие:

- анализ записей о подробных характеристиках вызовов. Назначение такого анализа — выявление категорий клиентов с похожими стереотипами пользования их услугами и разработка привлекательных наборов цен и услуг;
- выявление лояльности клиентов. Data Mining можно использовать для определения характеристик клиентов, которые, один раз воспользовавшись услугами данной компании, с большой долей вероятности останутся ей верными. В итоге средства, выделяемые на маркетинг, можно тратить там, где отдача больше всего.

Страхование. Страховые компании в течение ряда лет накапливают большие объемы данных. Здесь обширное поле деятельности для методов Data Mining:

- выявление мошенничества. Страховые компании могут снизить уровень мошенничества, отыскивая определенные стереотипы в заявлениях о выплате страхового возмещения, характеризующих взаимоотношения между юристами, врачами и заявителями.
- анализ риска. Путем выявления сочетаний факторов, связанных с оплаченными заявлениями, страховщики могут уменьшить свои потери по обязательствам. Известен случай, когда в США крупная страховая компания обнаружила, что суммы, выплаченные по заявлениям людей, состоящих в браке, вдвое превышает суммы по заявлениям одиноких людей. Компания отреагировала на это новое знание пересмотром своей общей политики предоставления скидок семейным клиентам.

7.4. Новые сферы применения Data Mining

- Медицина
- Молекулярная генетика и генная инженерия
- Прикладная химия

Медицина. Известно много экспертных систем для постановки медицинских диагнозов. Они построены главным образом на основе правил, описывающих сочетания различных симптомов различных заболеваний. С помощью таких правил узнают не только, чем болен пациент, но и как нужно его лечить. Правила помогают выбирать средства медикаментозного воздействия, определять показания — противопоказания, ориентироваться в лечебных процедурах, создавать условия наиболее эффективного лечения, предсказывать исходы назначенного курса лечения и т. п. Технологии Data Mining позволяют обнаруживать в медицинских данных шаблоны, составляющие основу указанных правил.

Молекулярная генетика и генная инженерия. Пожалуй, наиболее остро и вместе с тем четко задача обнаружения закономерностей в экспериментальных данных стоит в молекулярной генетике и геномной инженерии. Здесь она формулируется как определение так называемых маркеров, под которыми понимают генетические коды, контролирующие те или иные фенотипические признаки живого организма. Такие коды могут содержать сотни, тысячи и более связанных элементов. На развитие генетических исследований выделяются большие средства. В последнее время в данной области возник особый интерес к применению методов Data Mining. Известно несколько крупных фирм, специализирующихся на применении этих методов для расшифровки генома человека и растений.

Прикладная химия. Методы Data Mining находят широкое применение в прикладной химии (органической и неорганической). Здесь нередко возникает вопрос о выяснении особенностей химического строения тех или иных соединений, определяющих их свойства. Особенно актуальна такая задача при анализе сложных химических соединений, описание которых включает сотни и тысячи структурных элементов и их связей.

Data Mining применяется также во многих других областях:

- развитие автомобильной промышленности. При сборке автомобилей производители должны учитывать требования каждого отдельного клиента, поэтому им нужны возможность прогнозирования популярности определенных характеристик и знание того, какие характеристики обычно заказываются вместе;
- политика гарантий. Производителям нужно предсказывать число клиентов, которые подадут гарантийные заявки, и среднюю стоимость заявок;
- поощрение часто летающих клиентов. Авиакомпании могут обнаружить группу клиентов, которых данными поощрительными мерами можно побудить летать больше. Например, одна авиакомпания обнаружила категорию клиентов, которые совершали много полетов на короткие расстояния, не накапливая достаточно миль для вступления в их клубы, поэтому она таким образом изменила правила приема в клуб, чтобы поощрять число полетов так же, как и мили.

Можно привести еще много примеров различных областей знания, где методы Data Mining играют ведущую роль. Особенность этих областей заключается в их сложной системной организации. Они относятся главным образом к надкибернетическому уровню организации систем, закономерности которого не могут быть достаточно точно описаны на языке статистических или иных аналитических математических моделей. Данные в указанных областях неоднородны, гетерогенны, нестационарны и часто отличаются высокой размерностью.

7.5. Классификация систем Data Mining

Data Mining является мультидисциплинарной областью, возникшей и развивающейся на базе достижений прикладной статистики, распознавания образов, методов искусственного интеллекта, теории баз данных и др. Отсюда обилие методов и алгоритмов, реализованных в различных действующих системах Data Mining. Многие из таких систем интегрируют в себе сразу несколько подходов. Тем не менее, как правило, в каждой системе имеется какая-то ключевая компонента, на которую делается главная ставка. Выделенным классам дается краткая характеристика.

Предметно-ориентированные аналитические системы. Предметно-ориентированные аналитические системы очень разнообразны. Наиболее широкий подкласс таких систем, получивший распространение в области исследования финансовых рынков, носит название "технический анализ". Он представляет собой совокупность нескольких десятков методов прогноза динамики цен и выбора оптимальной структуры инвестиционного портфеля, основанных на различных эмпирических моделях динамики рынка. Эти методы часто используют несложный статистический аппарат, но максимально учитывают сложившуюся своей области специфику (профессиональный язык, системы различных индексов и пр.). На рынке имеется множество программ этого класса. Как правило, они довольно дешевы (обычно \$300–1000).

Статистические пакеты. Последние версии почти всех известных статистических пакетов включают наряду с традиционными статистическими методами также элементы Data Mining. Но основное внимание в них уделяется все же классическим методикам — корреляционному, регрессионному, факторному анализу и другим. Недостатком систем этого класса считают требование к специальной подготовке пользователя. Также отмечают, что мощные современные статистические пакеты являются слишком "тяжеловесными" для массового применения в финансах и бизнесе. К тому же часто эти системы весьма дороги — от \$1000 до \$15000. Есть еще более серьезный принципиальный недостаток статистических пакетов, ограничивающий их применение в Data Mining. Большинство методов, входящих в состав пакетов опираются на статистическую парадигму, в которой главными фигурантами служат усредненные характеристики выборки. А эти характеристики, как указывалось выше, при исследовании реальных сложных жизненных феноменов часто являются фиктивными величинами.

В качестве примеров наиболее мощных и распространенных статистических пакетов можно назвать SAS (компания SAS Institute), SPSS (SPSS), STATGRAPICS (Manugistics), STATISTICA, STADIA и другие.

Нейронные сети. Это большой класс систем, архитектура которых имеет аналогию (как теперь известно, довольно слабую) с построением нервной ткани из нейронов. В одной из наиболее распространенных архитектур, многослойном перцептроне с обратным распространением ошибки, имитируется работа нейронов в составе иерархической сети, где каждый нейрон более высокого уровня соединен своими входами с выходами нейронов нижележащего слоя. На нейроны самого нижнего слоя подаются значения входных параметров, на основе которых нужно принимать какие-то решения, прогнозировать развитие ситуации и т. д. Эти значения рассматриваются как сигналы, передающиеся в следующий слой, ослабляясь или усиливаясь в зависимости от числовых значений (весов), приписываемых межнейронным связям. В результате на выходе нейрона самого верхнего слоя вырабатывается некоторое значение, которое рассматривается как ответ — реакция всей сети на введенные значения входных параметров. Для того чтобы сеть можно было применять в дальнейшем, ее прежде надо "натренировать" на полученных ранее данных, для которых известны и значения входных параметров, и правильные ответы на них. Тренировка состоит в подборе весов межнейронных связей, обеспечивающих наибольшую близость ответов сети к известным правильным ответам.

Основным недостатком нейросетевой парадигмы является необходимость иметь очень большой объем обучающей выборки. Другой существенный недостаток заключается в том, что даже натренированная нейронная сеть представляет собой черный ящик. Знания, зафиксированные как веса нескольких сотен межнейронных связей, совершенно не поддаются анализу и интерпретации человеком (известные попытки дать интерпретацию структуре настроенной нейросети выглядят неубедительными – система “KINOsuite-PR”).

Примеры нейросетевых систем — BrainMaker (CSS), NeuroShell (Ward Systems Group), OWL (HyperLogic). Стоимость их довольно значительна: \$1500–8000.

Системы рассуждений на основе аналогичных случаев. Идея систем case based reasoning — CBR — на первый взгляд крайне проста. Для того чтобы сделать прогноз на будущее или выбрать правильное решение, эти системы находят в прошлом близкие аналоги наличной ситуации и выбирают тот же ответ, который был для них правильным. Поэтому этот метод еще называют методом "ближайшего соседа" (nearest neighbour). В последнее время распространение получил также термин memory based reasoning, который акцентирует внимание, что решение принимается на основании всей информации, накопленной в памяти. Системы CBR показывают неплохие результаты в самых разнообразных задачах. Главным их минусом считают то, что они вообще не создают каких-либо моделей или правил, обобщающих предыдущий опыт, — в выборе решения они основываются на всем массиве доступных исторических данных, поэтому невозможно сказать, на основе каких конкретно факторов CBR системы строят свои ответы. Другой минус заключается в произволе, который допускают системы CBR при выборе меры "близости". От этой меры самым решительным образом зависит объем множества прецедентов, которые нужно хранить в памяти для достижения удовлетворительной классификации или прогноза.

Примеры систем, использующих CBR, — KATE tools (Acknosoft, Франция), Pattern Recognition Workbench (Unica, США).

Деревья решений (decision trees). Деревья решения являются одним из наиболее популярных подходов к решению задач Data Mining. Они создают иерархическую структуру классифицирующих правил типа "ЕСЛИ... ТО..." (if-then), имеющую вид дерева. Для принятия решения, к какому классу отнести некоторый объект или ситуацию, требуется ответить на вопросы, стоящие в узлах этого дерева, начиная с его корня. Вопросы имеют вид "значение параметра А больше Х?". Если ответ положительный, осуществляется переход к правому узлу следующего уровня, если отрицательный — то к левому узлу; затем снова следует вопрос, связанный с соответствующим узлом. Популярность подхода связана как бы с наглядностью и понятностью. Но деревья решений принципиально не способны находить "лучшие" (наиболее полные и точные) правила в данных. Они реализуют наивный принцип последовательного просмотра признаков и "цепляют" фактически осколки настоящих закономерностей, создавая лишь иллюзию логического вывода. Вместе с тем, большинство систем используют именно этот метод. Самыми известными являются See5/C5.0 (RuleQuest, Австралия), Clementine (Integral Solutions, Великобритания), SIPINA (University of Lyon, Франция), IDIS (Information Discovery, США), KnowledgeSeeker (ANGOSS, Канада). Стоимость этих систем варьируется от 1 до 10 тыс. долл.

Эволюционное программирование. Проиллюстрируем современное состояние данного подхода на примере системы PolyAnalyst — отечественной разработке, получившей сегодня общее признание на рынке Data Mining. В данной системе гипотезы о виде зависимости целевой переменной от других переменных формулируются в виде программ на некотором внутреннем языке программирования. Процесс построения программ строится как эволюция в мире программ (этим подход немного похож на генетические алгоритмы). Когда система находит программу, более или менее удовлетворительно выражающую искомую зависимость, она начинает вносить в нее небольшие модификации и отбирает среди построенных дочерних программ те, которые повышают точность. Таким образом, система "выращивает" несколько генетических линий программ, которые конкурируют между собой в точности выражения искомой зависимости. Специальный модуль системы PolyAnalyst переводит най-

денные зависимости с внутреннего языка системы на понятный пользователю язык (математические формулы, таблицы и пр.). Другое направление эволюционного программирования связано с поиском зависимости целевых переменных от остальных в форме функций какого-то определенного вида. Например, в одном из наиболее удачных алгоритмов этого типа — методе группового учета аргументов (МГУА) зависимость ищут в форме полиномов. В настоящее время из продающихся в России систем МГУА реализован в системе NeuroShell компании Ward Systems Group. Стоимость систем до \$ 5000.

Генетические алгоритмы. Data Mining не основная область применения генетических алгоритмов. Их нужно рассматривать скорее как мощное средство решения разнообразных комбинаторных задач и задач оптимизации. Тем не менее генетические алгоритмы вошли сейчас в стандартный инструментарий методов Data Mining.

Первый шаг при построении генетических алгоритмов — это кодировка исходных логических закономерностей в базе данных, которые именуют хромосомами, а весь набор таких закономерностей называют популяцией хромосом. Далее для реализации концепции отбора вводится способ сопоставления различных хромосом. Популяция обрабатывается помощью процедур репродукции, изменчивости (мутаций), генетической композиции. Эти процедуры имитируют биологические процессы. Наиболее важные среди них: случайные мутации данных в индивидуальных хромосомах, переходы (кроссинговер) и рекомбинация генетического материала, содержащегося в индивидуальных родительских хромосомах (аналогично гетеросексуальной репродукции), и миграции генов. В ходе работы процедур на каждой стадии эволюции получают популяции с все более совершенными индивидуумами. Генетические алгоритмы удобны тем, что их легко распараллеливать. Например, можно разбить поколение на несколько групп и работать с каждой из них независимо, обмениваясь время от времени несколькими хромосомами. Существуют также и другие методы распараллеливания генетических алгоритмов.

Генетические алгоритмы имеют ряд недостатков. Критерий отбора хромосом и используемые процедуры являются эвристическими и далеко не гарантируют нахождения “лучшего” решения. Как и в реальной жизни, эволюцию может “заклинить” на какой-либо непродуктивной ветви. И, наоборот, можно привести примеры, как два неперспективных родителя, которые будут исключены из эволюции генетическим алгоритмом, оказываются способными произвести высокоэффективного потомка. Это особенно становится заметно при решении высокоразмерных задач со сложными внутренними связями. Примером может служить система GeneHunter фирмы Ward Systems Group. Его стоимость — около \$1000.

Алгоритмы ограниченного перебора. Алгоритмы ограниченного перебора были предложены в середине 60-х годов М.М. Бонгардом для поиска логических закономерностей в данных. С тех пор они продемонстрировали свою эффективность при решении множества задач из самых различных областей. Эти алгоритмы вычисляют частоты комбинаций простых логических событий в подгруппах данных. Примеры простых логических событий: $X = a$; $X < a$; $X > a$; $a < X < b$ и др., где X — какой-либо параметр, “a” и “b” — константы. Ограничением служит длина комбинации простых логических событий (у М. Бонгарда она была равна 3). На основании анализа вычисленных частот делается заключение о полезности той или иной комбинации для установления ассоциации в данных, для классификации, прогнозирования и пр.

Наиболее ярким современным представителем этого подхода является система WizWhy предприятия WizSoft. Хотя автор системы Абрахам Мейдан не раскрывает специфику алгоритма, положенного в основу работы WizWhy, по результатам тщательного тестирования системы были сделаны выводы о наличии здесь ограниченного перебора (изучались результаты, зависимости времени их получения от числа анализируемых параметров и др.). Автор WizWhy утверждает, что его система обнаруживает ВСЕ логические if-then правила в данных. На самом деле это, конечно, не так. Во-первых, максимальная длина комбинации в if-then правиле в системе WizWhy равна 6, и, во-вторых, с самого начала работы алгоритма производится эвристический поиск простых логических событий, на которых потом строит-

ся весь дальнейший анализ. Поняв эти особенности WizWhy, нетрудно было предложить простейшую тестовую задачу, которую система не смогла вообще решить. Другой момент — система выдает решение за приемлемое время только для сравнительно небольшой размерности данных.

Тем не менее, система WizWhy является на сегодняшний день одним из лидеров на рынке продуктов Data Mining. Это не лишено оснований. Система постоянно демонстрирует более высокие показатели при решении практических задач, чем все остальные алгоритмы. Стоимость системы около \$ 4000, количество продаж — 30000.

Системы для визуализации многомерных данных. В той или иной мере средства для графического отображения данных поддерживаются всеми системами Data Mining. Вместе с тем, весьма внушительную долю рынка занимают системы, специализирующиеся исключительно на этой функции. Примером здесь может служить программа DataMiner 3D словацкой фирмы Dimension5 (5-е измерение). В подобных системах основное внимание сконцентрировано на дружелюбности пользовательского интерфейса, позволяющего ассоциировать с анализируемыми показателями различные параметры диаграммы рассеивания объектов (записей) базы данных. К таким параметрам относятся цвет, форма, ориентация относительно собственной оси, размеры и другие свойства графических элементов изображения. Кроме того, системы визуализации данных снабжены удобными средствами для масштабирования и вращения изображений. Стоимость систем визуализации может достигать нескольких сотен долларов.

7.6. Развитие систем Data Mining

Рынок систем Data Mining экспоненциально развивается. В этом развитии принимают участие практически все крупнейшие корпорации (см., например, <http://www.kdnuggets.com>). В частности, Microsoft непосредственно руководит большим сектором данного рынка (издает специальный журнал, проводит конференции, разрабатывает собственные продукты).

Системы Data Mining применяются по двум основным направлениям: 1) как массовый продукт для бизнес-приложений; 2) как инструменты для проведения уникальных исследований (генетика, химия, медицина и пр.). В настоящее время стоимость массового продукта от \$1000 до \$10000. Количество инсталляций массовых продуктов, судя по имеющимся сведениям, сегодня достигает десятков тысяч. Лидеры Data Mining связывают будущее этих систем с использованием их в качестве интеллектуальных приложений, встроенных в корпоративные хранилища данных.

Несмотря на обилие методов Data Mining, приоритет постепенно все более смещается в сторону логических алгоритмов поиска в данных if-then правил. С их помощью решаются задачи прогнозирования, классификации, распознавания образов, сегментации БД, извлечения из данных “скрытых” знаний, интерпретации данных, установления ассоциаций в БД и др. Результаты таких алгоритмов эффективны и легко интерпретируются.

Вместе с тем, главной проблемой логических методов обнаружения закономерностей является проблема перебора вариантов за приемлемое время. Известные методы либо искусственно ограничивают такой перебор (алгоритмы KOPA, WizWhy), либо строят деревья решений (алгоритмы CART, CHAID, ID3, See5, Sipina и др.), имеющих принципиальные ограничения эффективности поиска if-then правил. Другие проблемы связаны с тем, что известные методы поиска логических правил не поддерживают функцию обобщения найденных правил и функцию поиска оптимальной композиции таких правил. Удачное решение указанных проблем может составить предмет новых конкурентоспособных разработок.

Вопросы для самоконтроля

1. Основные методы поиска и анализа данных?
2. Что означает термин Data Mining?
3. Классификация систем Data Mining?
4. Для каких целей применяют программное обеспечение систем Data Mining?
5. Особенности рынка компьютерного анализа данных?

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд. – Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд. – Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд., стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд. – Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 8

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА МОДЕЛИРОВАНИЯ НА ОСНОВЕ ДОКУМЕНТАЛЬНЫХ И ИНСТРУМЕНТАЛЬНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

8.1. Электронные документы

Документ является основным способом представления информации на любом современном предприятии. Неоспорима важность сохранности и умелого использования информационных ресурсов предприятия для успешного ведения бизнеса. Эффективность управления предприятием зависит и от того, насколько разумно в нем организовано управление документооборотом.

На любом предприятии, где ведется активная работа с различными документами, рано или поздно встает проблема систематизации, обработки и безопасного хранения значительных объемов информации. Важную роль в оптимизации деятельности предприятия любого размера и профиля деятельности играют современные системы электронного документооборота.

При этом под термином электронный документ в рассматриваемом контексте обычно понимается текстовый файл, документ MS Word, электронная таблица MS Excel, графический файл, несколько взаимосвязанных файлов разных форматов (так называемый составной документ) и др.

8.2. Современные системы электронного документооборота

Под управлением электронным документооборотом в общем случае принято понимать организацию движения документов между подразделениями предприятия, группами пользователей или пользователями. При этом под движением документов понимается не их физическое перемещение (т.к. они чаще всего остаются на сервере), а передачу прав на их использование с уведомлением конкретных пользователей и контролем за их исполнением. Главное назначение систем электронного документооборота это организация хранения электронных документов, а также работы с ними (в частности, их поиска, как по атрибутам, так и по содержанию). В системах электронного документооборота также реализован санкционированный доступ к документам, отслеживаются произведенные в них изменения и контролируются все их версии и подверсии.

В настоящее время на большинстве российских предприятий корпоративная информация (договоры, прайс-листы, приказы, деловая переписка, финансовая документация и др.) хранится как в бумажном, так и в электронном виде (часто неупорядоченно на файл-сервере). При этом практически неизбежно возникает дублирование информации, а также не обеспечивается поддержка разных версий документов и безопасность их хранения. При преобладании электронной формы хранения документов, они не всегда используются эффективно. Потребность в эффективном управлении электронными документами и привела к созданию систем электронного документооборота (первые из которых появились еще в 1980-е гг.). Эти системы еще называют Electronic Document Management Systems (EDMS).

EDMS-системы являются неотъемлемой частью мирового рынка программного обеспечения (ПО) управления электронным документооборотом и содержимым информационных систем, так называемого DCT (document and content technologies) рынка. Аналитики IDC выделяют следующие сегменты DCT-рынка: системы электронного документооборота; системы управления содержимым на порталах предприятий и системы управления содержимым для электронной коммерции. Рост потребности предприятий в повышении эффективности групповой работы с корпоративными документами неизбежно ведет к дальнейшему развитию данного рынка. По прогнозу IDC (отчет Document and Content Technologies Applications

Forecast and Analysis, 2000-2004), объем мирового DCT рынка должен увеличиться с \$1,1 млрд в 1999 г. до почти \$4,4 млрд в 2004 г. (с ежегодными темпами роста в 32%). Развитию рынка способствуют также распространение электронной коммерции и рост потребности предприятий в Web-совместимых интегрированных инструментальных средствах доступа к информации. При этом особенно быстро растет спрос потребителей на инструментальные средства сбора, поиска и анализа информации, с помощью которых возможно более оперативно обрабатывать разнородные собрания текстовых файлов, графических файлов, видео- и аудиофайлов. По оценке IDC в настоящее время продажи систем электронного документооборота формируют основу доходов на DCT-рынке.

8.3. Информационно-поисковые системы

Для быстрого поиска информации в Internet разработаны специальные программы, которые по заданным адресам и ссылкам мгновенно отыскивают нужную информацию. При этом число обработанных информационных ресурсов может достигать сотен тысяч.

Информационно-поисковая система (ИПС) – это Web-сайт, предоставляющий возможность поиска информации в Internet. Большинство поисковых систем ищут информацию на сайтах Всемирной паутины, но существуют также системы, способные искать файлы на Ftp-серверах, товары в Internet-магазинах, а также информацию в группах новостей Usenet.

В последнее время появился новый тип поисковых движков, основанных на технологии RSS, - семейство XML-форматов, предназначенных для описания лент новостей, анонсов статей, изменений в блогах и т. п. Можно назвать и аналогичные технологии: Rich Site Summary (стандарт RSS 0.9x) - обогащенная сводка сайта; RDF Site Summary (RSS 0.9 и 1.0) - сводка сайта с применением инфраструктуры описания ресурсов; Really Simple Syndication (RSS 2.x) - очень простое приобретение информации. Информация из различных источников, представленная в формате XML на базе RSS-стандартов, может быть собрана, обработана и представлена пользователю в удобном для него виде специальными программами-агрегаторами.

Комплекс программ, обеспечивающий функциональность поисковой системы, называют поисковым движком или поисковой машиной. Основными критериями качества работы поисковой машины являются релевантность, полнота базы, учет морфологии языка. Индексация информации осуществляется специальными поисковыми роботами. Улучшение работы поисковых систем - это одна из приоритетных задач сегодняшнего Internet.

Первой поисковой системой для Всемирной паутины был "Wandex" - робот, разработанный Мэтью Греем (Matthew Gray) из Массачусетского технологического института в 1993 году. В том же 1993 году появилась поисковая система Aliweb, работающая до сих пор. Первой полнотекстовой (т. н. "Crawler-based" - то есть индексирующей ресурсы при помощи робота) поисковой системой стала "WebCrawler", запущенная в 1994 году. В отличие от своих предшественников она позволяла пользователям искать по любым ключевым словам на любой Web-странице - с тех пор это стало стандартом во всех основных поисковых системах. Кроме того, это был первый поисковик, о котором стало известно в широких кругах. В 1994 году был запущен поисковик Lycos, разработанный в университете Карнеги Мелона (США).

Вскоре появилось множество других конкурирующих поисковых машин, таких как Excite, Infoseek, Inktomi, Northern Light и AltaVista. В некотором смысле они конкурировали с популярными Интернет-каталогами, такими как Yahoo!. Позже каталоги соединились или добавили к себе поисковые машины, чтобы увеличить функциональность. В 1996 году русскоязычным пользователям Internet стало доступно морфологическое расширение к поисковой машине AltaVista и оригинальные российские поисковые машины Rambler и Aport. 23 сентября 1997 года была реализована поисковая машина Yandex.

Помимо поисковых машин для Всемирной паутины существовали и поисковики для других протоколов, такие как Archie для поиска по анонимным FTP-серверам и Veronica для поиска в Gopher.

В настоящее время совокупности поисковых и сервисных программ образуют мощные общедоступные и коммерческие поисковые службы: в зарубежном секторе Internet это AltaVista, Excite, Google, HotBot, Infoseek (Go)Light, Lycos, Magellan, Northern, Yahoo!, Open Text, Web Crawler, в русскоязычном секторе основными полнотекстовыми поисковыми системами считаются Апорт, "Иван Сусанин", "Кирилл и Мефодий", "Россия-Он-Лайн", Rambler, List.ru, Russia on the Net, FTP-Search, Yandex.

По данным компании Net Applications, в декабре 2009 года рыночная доля Google в мире составляла 77,04%, Yahoo - 12,46%, MSN - 3,33%, Microsoft Live Search - 2,57%, AOL - 2,12%, Ask - 1,38%, AltaVista - 0,13%, Excite - 0,07%, Lycos - 0,02%, All the Web - 0,02%.

Популярные поисковые системы в русскоязычном сегменте Internet (Рунете):

- многоязычные: Google (18% Рунета), Yahoo! (1% Рунета) и принадлежащие этой компании поисковые машины: Overture, Inktomi, AltaVista, Alltheweb FAST-Engine, а также MSN (2% Рунета, принадлежит компании "Microsoft");
- русскоязычные: Апорт (1% Рунета), Rambler (18% Рунета), Yandex (47% Рунета), Mail.ru (7% Рунета), Webalta, Qwika, Gogo.ru, Turtle, Punto, Nigma, Darodag - поисковая система товаров, VisualWorld, Вершки Рунета - поиск по заглавным страницам.

Большинство русскоязычных поисковых систем индексируют и ищут тексты на многих языках - украинском, белорусском, английском и др. Отличаются же они от "всеязычных" систем, индексирующих все документы подряд, тем, что в основном индексируют ресурсы, расположенные в доменных зонах, где доминирует русский язык, или другими способами ограничивают своих роботов русскоязычными сайтами.

Наряду с универсальными поисковыми системами большой популярностью пользуются специализированные, такие как метапоисковые MetaCrawler.com и Nigma.ru, или осуществляющие "вертикальный" поиск (по конкретным типам: новости, картинки, видео, фото, вакансии, группы товаров и т. п.)

Вопросы для самоконтроля

1. Что называют электронным документом?
2. Что такое электронный документооборот?
3. Определение информационно-поисковой системы?
4. Наиболее популярные поисковые системы в русскоязычном сегменте Internet?
5. В чём специфика метапоисковых информационных систем?

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд. – Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 9

OLAP- МЕХАНИЗМЫ СОВРЕМЕННЫХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

9.1. Основные понятия

OLAP-технологии являются одной из наиболее эффективных методик анализа больших объемов данных. Краткое определение данной методике можно дать следующим образом.

OLAP (англ. OnLine Analytical Processing, аналитическая обработка в реальном времени) — технология организации хранения информации и доступа к ней, отвечающая принципам FASMI (Fast Analysis of Shared Multidimensional Information):

- Fast (Быстрый) — быстрый отклик системы, чтобы аналитик «не терял мысль» (не более 5 секунд),
- Analysis (Анализ) — полнофункциональный анализ максимально удобным способом,
- Shared (Разделяемый) — многопользовательский доступ, защита информации и т.п.,
- Multidimensional (Многомерный) — многомерное концептуальное представление данных в виде кубов с иерархическими измерениями,
- Information (Информация) — получение информации в нужном объеме, там, где она необходима.

Как видно из определения, требования к системе построения отчетности, которую действительно можно называть OLAP-системой, достаточно жесткие — за секунды отчеты должны строиться по миллионам записей, столь же быстро раскрываться по интересующим пользователя аналитикам, фильтроваться и перегруппировываться

9.2. Компьютерные системы поддержки принятия решений

Современные системы поддержки принятия решения (СППР) представляют собой системы, максимально приспособленные к решению задач повседневной управленческой деятельности, являются инструментом, призванным оказать помощь лицам, принимающим решения (ЛПР). С помощью СППР может производиться выбор решений некоторых неструктурированных и слабоструктурированных задач, в том числе и многокритериальных. СППР, как правило, являются результатом мультидисциплинарного исследования, включающего теории баз данных, искусственного интеллекта, интерактивных компьютерных систем, методов имитационного моделирования.

Ранние определения СППР (в начале 70-х годов прошлого века) отражали следующие три момента: (1) возможность оперировать с неструктурированными или слабоструктурированными задачами, в отличие от задач, с которыми имеет дело исследование операций; (2) интерактивные автоматизированные (то есть реализованные на базе компьютера) системы; (3) разделение данных и моделей. Приведем определения СППР: СППР — совокупность процедур по обработке данных и суждений, помогающих руководителю в принятии решений, основанная на использовании моделей.

СППР — это интерактивные автоматизированные системы, помогающие лицу, принимающему решения, использовать данные и модели для решения слабоструктурированных проблем.

СППР — это система, которая обеспечивает пользователям доступ к данным и/или моделям, так что они могут принимать лучшие решения.

Последнее определение не отражает участия компьютера в создании СППР, вопросы возможности включения нормативных моделей в состав СППР и др.

В настоящее время нет общепринятого определения СППР, поскольку конструкция СППР существенно зависит от вида задач, для решения которых она разрабатывается, от доступных данных, информации и знаний, а также от пользователей системы. Можно при-

вести, тем не менее, некоторые элементы и характеристики, общепризнанные, как части СППР:

СППР — в большинстве случаев — это интерактивная автоматизированная система, которая помогает пользователю (ЛПР) использовать данные и модели для идентификации и решения задач и принятия решений. Система должна обладать возможностью работать с интерактивными запросами с достаточно простым для изучения языком запросов.

9.3. Характеристики систем поддержки принятия решений

Современные СППР обладают следующими четырьмя основными характеристиками:

- СППР использует и данные, и модели;
- СППР предназначены для помощи менеджерам в принятии решений для слабоструктурированных и неструктурированных задач;
- они поддерживают, а не заменяют, выработку решений менеджерами;
- цель СППР — улучшение эффективности решений.

Идеальная СППР:

- оперирует со слабоструктурированными решениями;
- предназначена для ЛПР различного уровня;
- может быть адаптирована для группового и индивидуального использования;
- поддерживает как взаимозависимые, так и последовательные решения;
- поддерживает 3 фазы процесса решения: интеллектуальную часть, проектирование и выбор;
- поддерживает разнообразные стили и методы решения, что может быть полезно при решении задачи группой ЛПР;
- является гибкой и адаптируется к изменениям как организации, так и ее окружения;
- проста в использовании и модификации;
- улучшает эффективность процесса принятия решений;
- позволяет человеку управлять процессом принятия решений с помощью компьютера, а не наоборот;
- поддерживает эволюционное использование и легко адаптируется к изменяющимся требованиям;
- может быть легко построена, если может быть сформулирована логика конструкции СППР;
- поддерживает моделирование;
- позволяет использовать знания.

9.4. Классификация СППР

На уровне пользователя Haettenschwiler делит СППР на пассивные, активные и кооперативные СППР. Пассивной СППР называется система, которая помогает процессу принятия решения, но не может вынести предложение, какое решение принять. Активная СППР может сделать предложение, какое решение следует выбрать. Кооперативная позволяет ЛПР изменять, пополнять или улучшать решения, предлагаемые системой, посылая затем эти изменения в систему для проверки. Система изменяет, пополняет или улучшает эти решения и посылает их опять пользователю. Процесс продолжается до получения согласованного решения.

На концептуальном уровне Power (2003) [1] отличает СППР, управляемые сообщениями (Communication-Driven DSS), СППР, управляемые данными (Data-Driven DSS), СППР, управляемые документами (Document-Driven DSS), СППР, управляемые знаниями

(Knowledge-Driven DSS) и СППР, управляемые моделями (Model-Driven DSS). СППР, управляемые моделями, характеризуются в основном доступ и манипуляции с математическими моделями (статистическими, финансовыми, оптимизационными, имитационными). Отметим, что некоторые OLAP-системы, позволяющие осуществлять сложный анализ данных, могут быть отнесены к гибридным СППР, которые обеспечивают моделирование, поиск и обработку данных.

Управляемая сообщениями (Communication-Driven DSS) (ранее групповая СППР — GDSS) СППР поддерживает группу пользователей, работающих над выполнением общей задачи.

СППР, управляемые данными (Data-Driven DSS) или СППР, ориентированные на работу с данными (Data-oriented DSS) (также известные как Business Intelligence) в основном ориентируются на доступ и манипуляции с данными. СППР, управляемые документами (Document-Driven DSS), управляют, осуществляют поиск и манипулируют неструктурированной информацией, заданной в различных форматах. Наконец, СППР, управляемые знаниями (Knowledge-Driven DSS) обеспечивают решение задач в виде фактов, правил, процедур.

На техническом уровне Power (1997) [1] различает СППР всего предприятия и настольную СППР. СППР всего предприятия подключена к большим хранилищам информации и обслуживает многих менеджеров предприятия. Настольная СППР — это малая система, обслуживающая лишь один компьютер пользователя. Существуют и другие классификации (Alter [3], Holsapple и Whinston [2], Golden, Hevner и Power [1]). Отметим лишь, что превосходная для своего времени классификация Alter'a, которая разбивала все СППР на 7 классов, в настоящее время несколько устарела.

В зависимости от данных, с которыми эти системы работают, СППР условно можно разделить на оперативные и стратегические. Оперативные СППР предназначены для немедленного реагирования на изменения текущей ситуации в управлении финансово-хозяйственными процессами компании. Стратегические СППР ориентированы на анализ значительных объемов разнородной информации, собираемых из различных источников. Важнейшей целью этих СППР является поиск наиболее рациональных вариантов развития бизнеса компании с учетом влияния различных факторов, таких как конъюнктура целевых для компании рынков, изменения финансовых рынков и рынков капиталов, изменения в законодательстве и др. СППР первого типа получили название Информационных Систем Руководства (Executive Information Systems, ИСР). По сути, они представляют собой конечные наборы отчетов, построенные на основании данных из транзакционной информационной системы предприятия, в идеале адекватно отражающей в режиме реального времени основные аспекты производственной и финансовой деятельности. Для ИСР характерны следующие основные черты:

- отчеты, как правило, базируются на стандартных для организации запросах; число последних относительно невелико;
- ИСР представляет отчеты в максимально удобном виде, включающем, наряду с таблицами, деловую графику, мультимедийные возможности и т. п.;
- как правило, ИСР ориентированы на конкретный вертикальный рынок, например финансы, маркетинг, управление ресурсами.

СППР второго типа предполагают достаточно глубокую проработку данных, специально преобразованных так, чтобы их было удобно использовать в ходе процесса принятия решений. Неотъемлемым компонентом СППР этого уровня являются правила принятия решений, которые на основе агрегированных данных дают возможность менеджерам компании обосновывать свои решения, использовать факторы устойчивого роста бизнеса компании и снижать риски. СППР второго типа в последнее время активно развиваются. Технологии этого типа строятся на принципах многомерного представления и анализа данных (OLAP).

При создании СППР можно использовать Web-технологии. В настоящее время СППР на основе Web-технологий для ряда компаний являются синонимами СППР предприятия.

Архитектура СППР представляется разными авторами по-разному. Приведем пример. Marakas предложил обобщенную архитектуру, состоящую из 5 различных частей: (а) система управления данными (the data management system — DBMS), (b) система управления моделями (the model management system — MBMS), (c) машина знаний (the knowledge engine (KE)), (d) интерфейс пользователя (the user interface) и (e) пользователи (the user(s)).

Вопросы для самоконтроля

1. Определение информационно-поисковой системы?
2. Что называют поисковым движком или поисковой машиной?
3. Какова специфика поиска на основе RSS-стандартов?
4. Наиболее популярные поисковые системы в русскоязычном сегменте Internet?
5. В чём специфика метапоисковых информационных систем?

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд. – Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд. – Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд., стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд. – Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 10

ОСНОВЫ ТЕЛЕКОММУНИКАЦИЙ И СЕТЕВЫХ ТЕХНОЛОГИЙ

10.1. Общие принципы построения компьютерных сетей

По принципам построения, компьютерные сети делят на локальные и удаленные. Локальные сети создаются, как правило, в одной организации или в одном помещении.

Самым простым вариантом такой сети является связь компьютеров через параллельные или последовательные порты. В этом случае нет надобности в каком-либо дополнительном оборудовании. Необходимы только соединительные проводники. Такая связь между компьютерами настраивается в пределах одной комнаты. Используется она для передачи данных от одного компьютера к другому. В этом случае можно переносить данные без помощи магнитных, оптических дисков, Flash-памяти и т.п.. Любая оболочка современной операционной системы имеет программные средства, которые обеспечивают такую передачу данных.

В локальных одноранговых компьютерных сетях компьютеры подключаются к сети через специальные адаптеры сети, а функционирование сети поддерживается операционной системой сети. Примером таких операционных систем являются: Novell Personal Net Ware, Net Ware Line и др.

Все компьютеры и их операционные системы в локальных одноранговых компьютерных сетях должны быть однотипными. Пользователи этой сети могут передавать друг другу данные, использовать общие принтеры, магнитные и оптические диски и т.д.

В локальной многоранговой компьютерной сети используется один более мощный компьютер, который называется сервером, а другие, менее мощные - рабочими станциями. На серверах используется специальное системное обеспечение, которое отличается от системного программного обеспечения рабочих станций.

Удаленные компьютерные сети принято разделять на региональные и международные. Региональные создаются в определенных регионах, например, государстве, а международные обеспечивают связь любого компьютера с другим компьютером всемирной сети. Примером таких сетей является Relcom (для государств СНГ) и Internet (для всего мира). Из региональных компьютерных сетей можно выходить в Internet.

Связь компьютеров в региональных сетях обеспечивается обычными телефонными сетями или специально выделенными для этого сетями через специальные устройства, которые называются модемами. Модем преобразует сигналы двоичных кодов в звуковые сигналы языкового диапазона, и наоборот.

Компьютеры отдельного района (города) через модемы и линии связи подключаются к более мощному компьютеру, который называется провайдером. Каждому компьютеру пользователю, который подключен к сети, присваиваются реквизиты (адрес). Провайдеры, используя реквизиты, обеспечивают связь компьютеров отдельных пользователей.

Связь между компьютерами разных континентов осуществляется спутниковыми, радиорелейными или кабельными каналами связи.

Региональные компьютерные сети могут работать в разных режимах. Самый простой - режим электронной почты. Его используют для передачи с одного компьютера на другой писем, документов и т.д.

10.2. Локальные сети

Локальная сеть представляет собой набор компьютеров, периферийных устройств (принтеров и т. п.) и коммутационных устройств, соединенных кабелями. Локальные сети делятся на учрежденческие (офисные сети фирм, сети организационного управления и другие сети,

отличающиеся по терминологии, но практически одинаковые по своей структуре) и сети управления технологическими процессами на предприятиях.

Локальные сети характерны тем, что расстояния между компонентами сети сравнительно невелики, и, как правило, не превышают нескольких километров. Локальные сети различаются по роли и значению ЭВМ в сети, структуре, методам доступа пользователей к сети, способам передачи данных между компонентами сети и др. Каждой из предлагаемых на рынке сетей присущи свои достоинства и недостатки. Выбор сети определяется числом подключаемых пользователей, их приоритетом, необходимой скоростью и дальностью передачи данных, требуемой пропускной способностью, надежностью и стоимостью сети.

Основными видами локальных вычислительных сетей являются Ethernet и Arcnet. При этом Ethernet может иметь несколько типов кабеля:

- тонкий кабель Ethernet - иначе называется "Thinnet". Имеет ряд преимуществ, таких как использование более дешевого кабеля по сравнению с системой толстого кабеля Ethernet и использование аппаратуры, которую проще устанавливать;
- толстый кабель Ethernet (также известная как "Thicknet") получила свое название благодаря используемому в ней стандартному, или толстому кабелю Ethernet. Толстый кабель позволяет включать в систему большее количество компьютеров и увеличивать расстояние между компьютерами. Однако этот кабель дороже, а его установка сложнее по сравнению с тонким кабелем Ethernet;
- витая пара Ethernet. Преимущество системы Ethernet на витой паре в том, что кабель дешевле по сравнению с перечисленными выше кабелями, а его установка проще.

Наравне с приведенными выше способами подключения встречается способ Token-ring. Одним из преимуществ данной системы является прогнозируемость: одна часть сети может выйти из строя, но сеть в целом все-таки не остановится. Система поддерживается программным обеспечением для больших ЭВМ фирмы IBM, что может в некоторых ситуациях принести заметную выгоду. Слабые стороны системы в сравнении с другими системами заключаются в дороговизне и усложнении монтажа кабелей. К тому же, в некоторых случаях трудно вести поиск неисправностей.

Локальные сети, широко используемые в научных, управленческих, организационных и коммерческих технологиях, можно классифицировать по следующим признакам:

1. По роли ПЭВМ в сети:
 - сети с сервером;
 - одноранговые (равноправные) сети.
2. По структуре (топологии) сети:
 - одноузловые («звезда»);
 - кольцевые («кольцо»);
 - магистральные («шина»);
 - комбинированные.
3. По способу доступа пользователей к ресурсам и абонентам сети:
 - сети с подключением пользователя по указанным адресам абонентов по принципу коммутации каналов («звезда»);
 - сети с централизованным (программным) управлением подключения пользователей к сети («кольцо» и «шина»);
 - сети со случайной дисциплиной обслуживания пользователей («шина»).
4. По виду коммуникационной среды передачи информации:
 - сети с использованием существующих учреждений телефонных сетей;
 - сети на специально проложенных кабельных линиях связи;
 - комбинированные сети, совмещающие кабельные линии и радиоканалы.
5. По дисциплине обслуживания пользователей (способу доступа пользователей к сети):
 - приоритетные, задающиеся ЦУС, когда пользователи получают доступ к сети в соответствии с присвоенными им приоритетами (постоянными или изменяющимися);
 - неприоритетные, когда все пользователи сети имеют равные права доступа к сети.

6. По размещению данных в компонентах сети:
- с центральным банком данных;
 - с распределенным банком данных;
 - с комбинированной системой размещения данных.

10.3. Глобальные компьютерные сети

Глобальные компьютерные сети можно классифицировать по следующим признакам:

1. По типу средств коммуникаций:
 - наземные многоузловые сети;
 - спутниковые радиосети;
 - комбинированные сети.
2. По способу коммутации сообщений:
 - коммутация каналов;
 - коммутация сообщений;
 - коммутация пакетов;
 - адаптивная коммутация.
3. По выбору маршрута передачи сообщения:
 - фиксированные пути;
 - направленный выбор пути;
 - случайные пути;
 - лавинный способ.

Региональная сеть - это города, объединенные в сеть посредством расположенных в них компьютеров. К международной вычислительной сети следует отнести Internet. Это единственная сеть, объединяющая целые государства. На данный момент ведутся разработки по созданию альтернативной глобальной сети.

В России и СНГ в последние годы интенсивно внедряется сетевая компьютерная инфраструктура. Независимые государства развивают свои компьютерные сети и активно включаются в мировое информационное сообщество на базе глобальных международных сетей. В сетях России и СНГ основными каналами связи являются: коммутируемая телефонная сеть общего пользования, выделенные телефонные линии связи, специальные сети передачи данных (ПД-200, «Искра») и сеть абонентского телеграфа. В последнее время используются также линии связи на оптоволоконных кабелях, сотовая связь и радиорелейная связь.

Основные национальные сети, а также международные сети, услугами которых могут пользоваться граждане СНГ:

БЕЛИКОС - белорусский узел коммерческой сети СИТЕК, работающей на территории СНГ, Балтии и Болгарии.

ИКСМИР (информационно-коммерческая сеть «Мировой информационный рынок») - сеть функционирует в 12 регионах СНГ. Обеспечивает электронную почту, коммерческие предложения, рекламу, курсы валют, биржевые новости, цены на рынках, законодательства стран и расписания движения железнодорожного и авиационного транспорта.

СИТЕК - объединение национальных и региональных сетей коммерческого направления: биржевой и валютный рынки, товары и услуги, законодательство.

EUNET / RELCOM - международная коммерческая сеть, ориентированная в основном на предприятия и организации среднего класса. Популярность сети обусловлена приемлемым уровнем сервиса и относительно низкими ценами.

FIDONET - международная некоммерческая сеть, обеспечивающая свободный обмен информацией через BBS - электронные доски объявлений. Абоненты сети пользуются информацией BBS бесплатно.

РАУРАЛ - система электронных платежей, объединяет многие банки Беларуси, России, Украины, Казахстана и Кыргызстана, а также позволяет производить платежи в Азербайджане, Узбекистане и государствах Балтии.

SPRINTNET - крупнейшая в мире сеть электронной почты. Основной физической средой передачи данных является оптоволоконный кабель, включая трансатлантический канал. Сеть осуществляет передачу сообщений на факсимильные аппараты, средства телексной и телеграфной связи, обеспечивает электронные платежи и международные расчеты. Дает возможность пользователям доступа к большинству мировых сетей,

SWIFT - общество международных межбанковских финансовых телекоммуникаций. Сеть гарантирует оперативную пересылку и безопасное хранение финансовых документов абонентов в 130 странах мира и бесперебойное обслуживание клиентов в течение 24 часов.

Особая роль среди глобальных сетей принадлежит мировому сообществу сетей Internet. Internet - это глобальная компьютерная сеть, которая охватывает миллионы пользователей по всему миру. Эта сеть начала свою работу в 1969 году в качестве военного проекта. Она специально была спроектирована таким образом, чтобы остаться дееспособной даже в случае атомной войны. Информация, передаваемая через Internet, перемещается от одного компьютера к другому наиболее коротким путем. По этой причине, любые два компьютера останутся соединенными до тех пор, пока есть хотя бы один маршрут между ними. Люди узнают через Internet новости, слушают радио, смотрят на мир через разбросанные по всему миру видеокамеры, ходят в библиотеки и заочно обучаются в университетах других стран.

Вопросы для самоконтроля

1. В чём отличие локальных и удалённых компьютерных сетей?
2. Что такое одноранговые и многоранговые сети?
3. Какими каналами связи объединяются компьютеры разных континентов?
4. Что представляют собой локальные и глобальные компьютерные сети?
5. Как классифицируются компьютерные сети?

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>

2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>

3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд. – Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>

4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд. – Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>

2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16.

(обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>

3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд., стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>

4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд. – Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

Лекция 11

ВИДЫ УГРОЗ В ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ СИСТЕМАХ

11.1. Угрозы объектам информационной деятельности

Общая классификация угроз для автоматизированной информационной системы выглядит следующим образом:

Угрозы конфиденциальности данных и программ. Реализуются при несанкционированном доступе к данным (например, к сведениям о состоянии счетов клиентов банка), программам моделирования или каналам связи.

Информация, обрабатываемая на компьютерах или передаваемая по локальным сетям передачи данных, может быть несанкционированно снята через технические каналы утечки. При этом используется аппаратура, осуществляющая анализ электромагнитных излучений, возникающих при работе компьютера.

Такой съем информации представляет собой сложную техническую задачу и требует привлечения квалифицированных специалистов. С помощью приемного устройства, выполненного на базе стандартного телевизора, можно перехватывать информацию, выводимую на экраны дисплеев компьютеров с расстояния в тысячу и более метров. Определенные сведения о работе компьютерной системы извлекаются даже в том случае, когда ведется наблюдение за процессом обмена сообщениями без доступа к их содержанию.

Угрозы целостности данных, программ, аппаратуры. Целостность данных и программ нарушается при несанкционированном уничтожении, добавлении лишних элементов и модификации записей о состоянии счетов, изменении порядка расположения данных, формировании фальсифицированных платежных документов в ответ на законные запросы, при активной ретрансляции сообщений с их задержкой.

Несанкционированная модификация информации о безопасности системы может привести к несанкционированным действиям (неверной маршрутизации или утрате передаваемых данных) или искажению смысла передаваемых сообщений. Целостность аппаратуры нарушается при ее повреждении, похищении или незаконном изменении алгоритмов работы.

Угрозы доступности данных. Возникают в том случае, когда объект (пользователь или процесс) не получает доступа к законно выделенным ему службам или ресурсам. Эта угроза реализуется захватом всех ресурсов, блокированием линий связи несанкционированным объектом в результате передачи по ним своей информации или исключением необходимой системной информации.

Эта угроза может привести к ненадежности или плохому качеству обслуживания в системе и, следовательно, потенциально будет влиять на достоверность и своевременность доставки платежных документов.

Угрозы отказа от выполнения транзакций. Возникают в том случае, когда легальный пользователь передает или принимает платежные документы, а потом отрицает это, чтобы снять с себя ответственность.

Оценка уязвимости автоматизированной информационной системы и построение модели воздействий предполагают изучение всех вариантов реализации перечисленных выше угроз и выявление последствий, к которым они приводят.

Следует отметить, что задачу контроля доступа при одновременном обеспечении целостности ресурсов надежно решает только шифрование информации.

11.2. Компьютерные вирусы и антивирусные программы защиты

Компьютерный вирус - это некоторая специальная программа, выполняющая на компьютере пользователя всевозможные вредные действия и, как правило, запускающаяся без его ведома. При этом в зависимости от типа вируса может происходить порча или несанкционированная передача посторонним лицам информации, хранящейся в памяти компьютера, исчезновение информации или программ, существенное замедление работы компьютера, некорректная работа некоторых приложений, появление посторонних видео-аудио эффектов, остановка работы компьютера (т.е. его «зависание» или невозможность загрузки). Заражение вирусами происходит при использовании инфицированных носителей информации или через компьютерную сеть.

Цикл существования любого вируса состоит из трех основных этапов.

1. Неактивный цикл. В этом состоянии вирус внедрен в тело какого-либо файла и «ждет» своего часа. Именно в неактивном состоянии вирусы переносятся вместе с программами с одного компьютера на другой.

2. Активизация вируса происходит при запуске инфицированной программы или файла или при использовании инфицированного оборудования. При запуске вирус может создать в памяти компьютера резидентную копию или немедленно произвести вредоносные действия, запрограммированные в нем (т. е. перейти к третьему этапу). Если создана резидентная копия, то до момента, когда вирус себя обнаружит, может пройти достаточно большой «инкубационный период» (в зависимости от настроек, заложенных в вирус создателями). При активизации резидентной копии происходит размножение вируса, т. е. внедрение его кода в код других программ.

3. Внешние проявления «жизнедеятельности» вируса.

Защиту от существующих вирусов осуществляют антивирусные программы защиты, которые производят диагностику имеющихся на компьютере файлов и вновь поступающей информации, лечат или удаляют вирусы, помещают подозрительные файлы в «карантин».

В настоящее время на ПК наибольшее распространение получили следующие антивирусные программы: NOD32, Doctor Web, Symantec Antivirus, Kaspersky. Все современные программы защиты от вирусов, тем не менее, требуют периодического обновления антивирусных баз с целью отслеживания и идентификации вновь создаваемых вирусов

11.3. Классификация антивирусных программ

Таблица 11.1 - Типы антивирусных программ

Наименование	Описание
Программы - мониторы	Принцип работы состоит в отслеживании работы всех запускаемых или уже запущенных программ
Программы - сканеры	Запуск программы активизирует последовательный поиск программных кодов вирусов, возможно хранящихся в памяти компьютера и на носителях информации (винчестерах, дискетах, Flash-памяти)
Программы - фильтры	Работа данного рода программ направлена на фильтрацию данных, поступающих на компьютер по сети, в том числе происходит контроль над входящими электронными письмами и открываемыми web-страницами. К таким программам можно отнести программы - брандмауэры, контролирующие входящую и исходящую с компьютера информацию. В некоторых случаях реализация подобного рода защиты от вирусов возможна при помощи аппаратных средств
Программы - детекторы программного обеспечения	Осуществляют контроль за возможной работой «лишних» программ, запускающихся автоматически вместе с другим программным обеспечением, что может значительно замедлить работу ЭВМ

Вопросы для самоконтроля

1. Какие виды угроз информационным объектам в настоящее время получили наибольшее распространение?
2. Самое надёжное средство защиты информационных ресурсов?
3. Что называют компьютерным вирусом?
4. Циклы существования вируса?
5. Каковы основные типы антивирусных программ?

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

УСТАНОВКА, НАСТРОЙКА И ЭКСПЛУАТАЦИЯ АНТИВИРУСНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

12.1. Функции современного антивирусного программного обеспечения

Для организации эффективной антивирусной защиты необходимо наличие соответствующего программного обеспечения. Несмотря на все разнообразие современных антивирусных программных продуктов, принципы их работы одинаковы. К основным функциям современных антивирусов относятся:

- сканирование памяти и содержимого дисков по расписанию;
- сканирование памяти компьютера, а также записываемых и читаемых файлов в реальном режиме времени с помощью резидентного модуля;
- выборочное сканирование файлов с измененными атрибутами (размером, датой модификации, контрольной суммой и т. д.);
- сканирование архивных файлов;
- распознавание поведения, характерного для компьютерных вирусов;
- удаленная установка, настройка и администрирование антивирусных программ с консоли системного администратора; оповещение системного администратора о событиях, связанных с вирусными атаками, по электронной почте, пейджеру и т. д.;
- принудительная проверка подключенных к корпоративной сети компьютеров, инициируемая системным администратором;
- удаленное обновление антивирусного программного обеспечения и баз данных с информацией о вирусах, в том числе автоматическое обновление баз данных по вирусам посредством Internet;
- фильтрация трафика Internet на предмет выявления вирусов в программах и документах, передаваемых посредством протоколов SMTP, FTP, HTTP;
- выявление потенциально опасных Java-апплетов и модулей ActiveX;
- функционирование на различных серверных и клиентских платформах, а также в гетерогенных корпоративных сетях;
- ведение протоколов, содержащих информацию о событиях, касающихся антивирусной защиты.

В связи с тем, что одной из основных характеристик современных вирусных атак является их высокая скорость распространения и высокая частота появления новых атак, современное антивирусное программное обеспечение нужно обновлять как можно чаще, тем самым повышая качество защиты. Необходимо учитывать все актуальные на текущий момент времени вирусные угрозы. Но наличие антивирусного программного обеспечения — это обязательное, но не достаточное условие для отражения вирусной атаки. Мало иметь в своем распоряжении средство, следует продумать и методы его правильного использования. Защита от вирусов должна быть элементом политики безопасности, которую понимают и соблюдают все пользователи системы.

В настоящее время обычная корпоративная компьютерная сеть отечественного заказчика включает в себя десятки и сотни рабочих станций, десятки серверов, разнообразное активное и пассивное телекоммуникационное оборудование и, как правило, имеет очень сложную структуру.

Стоимость обслуживания такой сети катастрофически растет вместе с ростом числа подключенных рабочих станций. Сейчас все только и говорят о том, как в данных условиях можно уменьшить совокупную стоимость владения или эксплуатации компьютерной инфраструктуры предприятия. Очевидно, что расходы на антивирусную защиту корпоративной сети здесь являются не последним пунктом в списке общих расходов предприятия. Однако

существует принципиальная возможность оптимизации и снижения этих расходов путем использования специальных решений, позволяющих централизованно управлять антивирусной защитой корпоративной сети в реальном масштабе времени. Необходимо, чтобы такие решения позволяли администраторам сети предприятия отслеживать все точки проникновения вирусов с единой консоли управления и согласно технологии «клиент-сервер» эффективно управлять всеми присутствующими в корпоративной сети антивирусными средствами различных производителей.

Такая стратегия антивирусной защиты позволяет блокировать все возможные точки проникновения вирусов, такие как:

- проникновение вирусов на рабочие станции при использовании на рабочей станции инфицированных файлов с переносимых источников (флоппи-диски, компакт-диски, Zip, Jazz, Floptical и т. д.);
- заражение вирусами с помощью бесплатного инфицированного программного обеспечения, полученного из Internet через Web или FTP и сохраненного на локальной рабочей станции;
- проникновение вирусов при подключении к корпоративной сети инфицированных рабочих станций удаленных или мобильных пользователей;
- заражение вирусами с удаленного сервера, подсоединенного к корпоративной сети и обменивающегося инфицированными данными с корпоративными серверами приложений и баз данных;
- распространение электронной почты, содержащей в приложениях файлы Excel и Word, инфицированные макровирусами.

Однако именно требование комплексного централизованного управления стало камнем преткновения для успешного создания эффективных комплексных систем антивирусной защиты корпоративных сетей в отечественных компаниях, что в конечном счете привело к столь широкому проникновению компьютерных вирусов в сети Internet/intranet. Использование локальных антивирусных решений в корпоративной сети необходимо, но не достаточно для эффективной реализации антивирусной защиты предприятия. Сложившаяся сегодня ситуация требует незамедлительного вмешательства соответствующих должностных лиц и принятия решений, направленных на обеспечение и создание систем антивирусной защиты предприятия.

12.2. Методы предотвращения вирусных атак

Лучший способ борьбы с вирусной атакой - ее предотвращение. Для решения этой задачи необходимо:

- соответствующим образом сконфигурировать антивирусное программное обеспечение;
- использовать только лицензионное программное обеспечение;
- ограничить набор программ, которые пользователь способен установить в системе;
- устранить известные уязвимости в используемом программном обеспечении;
- контролировать использование накопителей гибких дисков и дисков CD-ROM; G разработать политику обработки электронной почты;
- разработать политику безопасности приложений, обрабатывающих документы с интерпретируемыми языками.
- чтобы соответствующим образом сконфигурировать антивирусное программное обеспечение, необходимо произвести следующие установки антивируса:
- сканирование в режиме реального времени, в фоновом или аналогичном режиме, должно быть разрешено;
- при старте системы нужно сканировать память, загрузочный сектор и системные файлы;
- своевременно обновлять вирусные базы данных;

- желательно сканировать файлы всех типов или, как минимум, COM-, EXE-файлы, а также файлы типа VBS, SHS, OCX;
- установить аудит всех действий антивирусных программ.

12.3. Источники вирусной опасности

Поскольку программное обеспечение, полученное из неизвестного источника, может быть троянским или зараженным вирусом, то необходимо пользоваться только лицензионным программным обеспечением.

Ограничение набора программ, которые пользователь способен установить в системе, связано с тем, что эти программы могут быть заражены вирусами или служить причиной успеха других атак. Особо следует обратить внимание на различные сервисы Internet и, в первую очередь, на программы передачи сообщений, такие как IRC, ICQ, Microsoft Chat (они могут передавать файлы и служить источником заражения системы).

Для устранения известных «дыр» в используемом программном обеспечении в качестве источника информации об уязвимостях можно использовать базы данных, которые обычно публикуются в списках рассылки Internet, а также на специальных сайтах.

Вся информация, содержащаяся на гибких и компакт-дисках, должна быть проверена на наличие вирусов до того, как с ней будут работать пользователи компьютерной системы.

В связи с тем, что сообщения электронной почты — один из самых популярных и быстрых способов распространения вирусов, в каждой организации должна быть разработана политика обработки электронной почты. Для защиты от проникновения вирусов через сообщения электронной почты каждый пользователь системы должен:

- никогда не открывать сразу почтовое вложение в пришедшем ему сообщении, а сохранять его в определенном «карантинном» каталоге; никогда не открывать почтовых вложений, которые не были запрошены или о которых не было уведомления от отправителя (даже когда отправитель известен, сообщение может содержать вирус, если отправитель неизвестен, сообщение с вложением лучше всего удалить); перед открытием вложения обязательно проверить его с помощью антивирусного программного обеспечения;
- если после выполнения всех этих процедур остались сомнения, стоит связаться с отправителем и выяснить у него информацию о посланном вложении;
- устранить возможные уязвимости в клиентском почтовом программном обеспечении.

Если пользователь или организация используют приложения, обрабатывающие документы с интерпретируемыми языками (например, семейство продуктов Microsoft Office), то порядок работы с этими документами тоже должен быть отражен в политике безопасности.

12.4. Новые типы антивирусного программного обеспечения

Обычно анализ вирусов заключается в выделении в них сигнатур и последующем их поиске в потенциальных объектах вирусной атаки. Таким образом, еще несколько лет назад достаточно было поймать вирус, изучить его код (для профессионалов это, как правило, было делом нескольких минут) и выделить сигнатуру. Но вирусные технологии не стояли на месте. Разрабатывались новые вирусы, а вслед за ними и новые программные антивирусные продукты.

Антивирусных средств довольно много. А так как в каждом конкретном случае надо выбирать антивирусный комплект, исходя из общей концепции информационной безопасности организации и нужд конкретного пользователя, то ниже кратко описаны основные типы антивирусных средств.

Существуют следующие стандартные программы защиты:

- детекторы (scanner);
- фаги (полифаги) (scanner/cleaner, scanner/remover);
- ревизоры;

- сторожа;
- специальные вакцины;
- блокировщики.

В большинстве случаев вирус, заразивший компьютер, помогут обнаружить уже разработанные программы-детекторы. Они проверяют, имеется ли в файлах на указанном пользователем диске специфическая для данного вируса последовательность байт. При обнаружении вируса программа выводит на экран соответствующее сообщение. Назначение детектора - только обнаружить вирус. Борьбаться с ним предстоит либо другой антивирусной программе, либо системному программисту.

Среди детекторов можно выделить эвристические анализаторы кода — набор подпрограмм, анализирующих код исполняемых файлов, памяти или загрузочных секторов для обнаружения в нем разных типов компьютерных вирусов. Обычно применяют универсальную схему кодоанализатора. Действуя в соответствии с этой схемой, кодоанализатор способен максимально эффективно использовать всю информацию, собранную для тестируемого объекта.

12.5. Эвристический метод распознавания новых вирусов

Эвристический подход состоит в попытке предложить, может быть, далекое от оптимального, но быстрое, решение для чрезвычайно сложных (или даже неразрешимых) проблем на основе все более и более достоверных предположений.

Основная идея такого подхода состоит в том, что эвристика сначала рассматривает поведение программы, а затем сопоставляет его с характерным для злонамеренной атаки, наподобие поведения троянского коня. Установить модель поведения и принять решение относительно него можно с помощью нескольких механизмов. Для того чтобы выявить и определить все возможные действия программы, используют два подхода: сканирование и эмуляция.

Подход со сканированием предполагает поиск «поведенческих штампов», например, наиболее типичных низкоуровневых способов открытия файлов. Или процедура сканирования обычного исполняемого файла просматривает все места, где программа открывает другой файл, и определяет, какого рода файлы она открывает и что в них записывает.

Второй метод определения поведения — эмуляция. Такой подход несколько сложнее. Программа пропускается через эмулятор Windows или макроэмулятор Macintosh или Word с целью посмотреть, что она будет делать. Однако возникают вопросы, потому что в этом случае многое зависит от причуд вирусов. Например, если вирус запрограммирован на форматирование вашего жесткого диска 25 февраля в 10 час. утра, а при эмуляции этого вируса на симуляторе дата установлена на 24 февраля, то вирус пока не проявит свои намерения.

Эффективность быстрого распознавания состоит в сочетании двух подходов и получении наиболее подробного каталога поведенческих штампов за возможно более короткое время. Для проверки факта заражения файла вирусом специалисты могут использовать различные варианты искусственного интеллекта — экспертные системы и нейронные сети.

Недостаток эвристического подхода состоит как раз в его эвристичности. Всегда есть вероятность, что чрезвычайно подозрительный файл в действительности совершенно безобиден. Однако последний эвристический механизм Symantec под названием Bloodhound позволяет обнаружить до 80% неизвестных вирусов выполняемых файлов и до 90% неизвестных макровирусов.

Стоит также заметить, что программы-детекторы не слишком универсальны, поскольку способны обнаружить только известные вирусы. Некоторым таким программам можно сообщить специальную последовательность байт, характерную для какого-то вируса, и они смогут обнаружить инфицированные им файлы: например, это умеют NotronAntiVirus или AVP-сканер.

12.6. Особенности применения антивирусных программ

Наиболее широкое распространение получили программы DrWeb и Kaspersky AVP. Благодаря своим новейшим детекторам, они могут обнаружить любые вирусы: как самые старые, так и только что появившиеся. Еще нужно упомянуть детектор ADInf. Эта антивирусная программа обнаруживает все вирусы, не изменяющие длину файлов, невидимые вирусы, и многие другие. Таким образом, эти три программы обеспечивают мощнейшую защиту против вирусов. Все эти программы можно вписать в файл AUTOEXEC.BAT, тогда при загрузке компьютера проверка на заражение вирусом будет проводиться автоматически. Кстати, на Западе тоже предпочитают пользоваться такими российскими программами, как DrWeb и AVP.

Несколько лет назад детекторы практически уступили свои позиции программам, называемых полифагами, но сегодня они вновь возвращаются на компьютерный рынок.

Для тех, кто пользуется только лицензионным программным обеспечением, нет необходимости тратить время на лечение зараженных вирусом файлов. Проще восстановить зараженную программу с дистрибутива. Но в связи с тем, что даже во многих достаточно крупных организациях очень часто используют не лицензионную, а «пиратскую» продукцию (возможно, уже зараженную вирусом), то и чистые детекторы (сканеры) еще не скоро будут в состоянии конкурировать с фагами. :

Фаги (полифаги) (scanner/cleaner, scanner/remover) - программы, способные не только обнаруживать, но и уничтожать вирусы, т. е. лечить «больные» программы (полифаг может уничтожить много вирусов). К полифагам относится и такая старая программа, как Aidtest, которая обнаруживает и обезвреживает около 2000 вирусов.

Основной принцип работы традиционного фага прост и секретом не является. Для каждого вируса путем анализа его кода, способов заражения файлов и т. д. выделяется некоторая характерная только для него последовательность байт. Эта последовательность называется сигнатурой данного вируса. Поиск вирусов в простейшем случае сводится к поиску их сигнатур (так работает любой детектор). Современные фаги используют другие методы поиска вирусов.

После обнаружения вируса в телепрограммы (или загрузочного сектора, который тоже, впрочем, содержит программу начальной загрузки) фаг обезвреживает его. Для этого разработчики антивирусных средств тщательно изучают работу каждого конкретного вируса: что он портит, как он портит, где он прячет то, что испортит (если прячет, конечно). В большинстве случаев фаг в состоянии благополучно удалить вирус и восстановить работоспособность испорченных программ. Но необходимо хорошо понимать, что это возможно далеко не всегда.

Программы, называемые ревизорами, контролируют возможные пути распространения инфекции. Изобретательность авторов вредоносного программного обеспечения ограничена некоторыми рамками, исходя из того, возможно в принципе. Эти рамки хорошо известны и поэтому вирусы все же не всемогущи. Если взять под контроль все мыслимые направления вирусной атаки на компьютер, то можно находиться практически в полной безопасности. Из программ-ревизоров, которые можно приобрести в России, следует обратить внимание на уже упомянутую выше программу ADInf.

Сторожами называют небольшие резидентные программы, находящиеся постоянно в памяти компьютера и контролирующие операции, которые они считают подозрительными. В качестве примера программы-сторожа можно привести известный программный продукт VSAFE.

Поскольку и вирусы, и обычные программы выполняют одни и те же операции, то невозможно даже выделить класс исключительно «вирусных» операций. Вследствие этого сторож либо вынужден ничего не контролировать и пассивно наблюдать за происходящим, либо «звенеть» при каждой подозрительной операции. Поэтому целесообразно использовать программы-сторожа на самом минимальном уровне контроля (например, отслеживания из-

менений загрузочных секторов). Такими сторожевыми функциями обладают некоторые современные BIOS, хотя и с этим все не так просто. Эта функция BIOS может конфликтовать с некоторыми операционными системами, а иногда может вообще не работать.

Специальные вакцины предназначены для обработки файлов и загрузочных секторов. Вакцины бывают пассивными и активными. Активная вакцина, «заражая» файл, подобно вирусу, предохраняет его от любого изменения и в ряде случаев способна не только обнаружить сам факт заражения, но и вылечить файл. Пассивные вакцины применяют только для предотвращения заражения файлов некоторыми вирусами, использующими простые признаки их зараженности — «странные» время или дата создания, определенные символьные строки и т. д.

В настоящее время вакцинирование широко не применяется. Бездумное вакцинирование всего и вся способно вызвать целые эпидемии несуществующих вирусных болезней. Так, в течение нескольких лет на территории бывшего СССР свирепствовала эпидемия вируса TIME. Жертвой этого вируса стали сотни абсолютно здоровых программ, обработанных антивирусной программой ANTI-kod.

Еще одной разновидностью антивирусных программ являются блокировщики вируса. Они позволяют ограничить распространение эпидемии, пока вирус не будет уничтожен. Практически все резидентные вирусы определяют факт своего присутствия в памяти машины, вызывая какое-либо программное прерывание с «хитрыми» параметрами. Если написать простую резидентную программу, которая будет имитировать наличие вируса в памяти компьютера, правильно «отзываясь» на определенный пароль, то вирус, скорее всего, сочтет эту машину уже зараженной.

Даже если некоторые файлы на компьютере содержат в себе код вируса, при использовании блокировщика заражения всех остальных файлов не произойдет. Для нормальной работы такой программы необходимо запустить блокировщик раньше всех остальных программ, например, в файле CONFIG.SYS. Но если вирус успел заразить COMMAND.COM или стартует из загрузочного сектора, то антивирус-блокировщик не поможет.

Очень важно применять альтернативные антивирусные решения. Сами по себе антивирусные сканеры и настройки защиты в различных приложениях не обеспечивают адекватной защиты от вредоносных программ. Антивирусные сканеры необходимо постоянно обновлять, хотя быстро распространяющиеся вирусы могут опередить эти модернизации.

Единственный способ избежать воздействия вредоносного программного обеспечения — блокировать подозрительные файлы на межсетевом экране или на шлюзе электронной почты. Многие организации сейчас блокируют все входящие присоединенные файлы, имеющие следующие, потенциально опасные, расширения: EXE, CORN, SCR, HTA, HTO, ASF, CHM, SHS, PIF. Другие устанавливают еще более жесткие фильтры, блокируя файлы с расширениями ADE, ADP, BAS, BAT, CMD, CNF, CPL, CRT, CSS, HIP, INF, INS, ISP, JS, JSE, INK, MDB, MDE, MSC, MSI, MSP, MST, PCD, REG, SET, SHB, URI, VB, VBE, VBS, WSC, WSF, WSH.

Один из ключевых вопросов, который будет стоять перед индустрией детектирующих систем в течение ближайших лет, заключается в том, продолжат ли заказчики покупать эти системы как самостоятельные продукты или уже скоро они станут приобретать их в комплекте с сетевым оборудованием — маршрутизаторами, коммутаторами или устройствами для локальных сетей. Ответ на него пока не найден, однако не приходится сомневаться, что системы выявления сетевых атак, используемые сегодня преимущественно такими крупными организациями, как банки и федеральные учреждения, со временем найдут дорогу к более широкому слою корпоративных пользователей.

Вопросы для самоконтроля

1. Основные функции антивирусного программного обеспечения?
2. Современная стратегия антивирусной защиты информации?
3. Методы предотвращения вирусных атак?
4. Основные источники вирусной опасности?
5. Идея эвристического подхода к распознаванию новых вирусов?

СПИСОК ЛИТЕРАТУРЫ

Основная (ЭСБ)

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>

Дополнительная (ЭБС)

1. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
2. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
3. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
4. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. **Советов, Б.Я., Цехановский, В.В.** Информационные технологии: теоретические основы: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 448 с. ISBN 978-5-8114-1912-8. Режим доступа: <http://elanbook.com/book/71733>
2. **Белов, В.В.** Алгоритмы и структуры данных: Учебник / Белов В.В., Чистякова В.И. - М.: КУРС, НИЦ ИНФРА-М, 2017. - 240 с.: 60х90 1/16. - (Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-906818-25-6 Режим доступа: <http://znanium.com/bookread2.php?book=766771>
3. **Петров, А.В.** Моделирование процессов и систем: учебное пособие. – 1-е изд.– Изд-во «Лань», 2015. – 288 с. ISBN 978-5-8114-1886-2. Режим доступа: <http://e.lanbook.com/book/68472>
4. **Бурнаева, Э.Г., Леора, С.Н.** Обработка и представление данных в MS Excel: учебное пособие. – 1-е изд. – Изд-во «Лань», 2016. – 160 с. ISBN 978-5-8114-1923-4. Режим доступа: <http://elanbook.com/book/71706>
5. **Попов, А.М., Сотников, В.М., Нагаева, В.И.** Информатика и математика: учебное пособие. – 1-е изд.– Изд-во «ЮНИТИ-ДАНА», 2012. – 302 с. ISBN 978-5-238-01396-1. Режим доступа: <http://www.iprbookshop.ru/7039.html>
6. **Радаева, Я.Г.** Word 2010: Способы и методы создания профессионально оформленных документов: Учебное пособие / Я.Г. Радаева. - М.: Форум: НИЦ ИНФРА-М, 2013. - 160 с.: 70х100 1/16. (обложка) ISBN 978-5-91134-736-9, 500. Режим доступа: <http://znanium.com/bookread2.php?book=402060>
7. **Нестеров, А.С.** Основы информационной безопасности: учебное пособие. – 2-е изд.. стер. – Изд-во «Лань», 2016. – 324 с. ISBN 978-5-8114-2290-6. Режим доступа: <http://elanbook.com/book/75515>
8. **Васильев, А.Н.** Числовые расчеты в Excel: справочник. – 1-е изд.– Изд-во «Лань», 2014. – 608 с. ISBN 978-5-8114-1580-9. Режим доступа: <http://e.lanbook.com/book/68464>

СОДЕРЖАНИЕ

Введение	3
Лекция 1. Основные понятия теории систем	4
1.1. Системный подход. Основные понятия	4
1.2. Определение системы	4
1.3. Основные свойства системы	6
1.3. Система и внешняя среда. Входные и выходные величины	7
Вопросы для самоконтроля	9
Список литературы	9
Лекция 2. Системный анализ	10
2.1. Классификация систем. Виды систем	10
2.2. Структура систем. Иерархические структуры	10
2.3. Состояние и движение системы	12
2.4. Методы изучения состояния и движения системы	14
2.5. Типы поведения динамических систем. Понятие устойчивости	16
2.6. Кибернетические системы	17
2.7. Особенности систем в сфере пищевых производств	18
Вопросы для самоконтроля	20
Список литературы	20
Лекция 3. Математическое моделирования на основе информационных технологий	22
3.1. Этапы решения практических задач на ЭВМ	22
3.2. Основные понятия моделирования	22
3.3. Общая классификация моделей	23
3.4. Классификация математических моделей	23
3.5. Оптимизационные математические модели	24
3.6. Современные методы компьютерного моделирования	25
3.7. Структурно-функциональное моделирование	26
Вопросы для самоконтроля	27
Список литературы	28
Лекция 4. Современные методы математического программирования и моделирования	29
4.1. Вводные замечания	29
4.2. Линейное математическое программирование	29
4.3. Методы решения задач линейного программирования	31
4.4. Двойственная задача линейного программирования	32
4.5. Двойственный симплекс-метод	33
4.6. Целочисленное математическое программирование	33
4.7. Области применения и ограничения линейного программирования	34
4.8. Методы построения опорных планов	35
Вопросы для самоконтроля	36
Список литературы	37
Лекция 5. Вероятностные математические модели	38
5.1. Основные определения	38
5.2. Числовые характеристики случайных величин	39
5.3. Вычисление числовых характеристик случайных величин	40
5.4. Непрерывные случайные величины	41
5.5. Модельные распределения вероятностей	43

5.6. Равномерное распределение	44
5.7. Экспоненциальное (показательное) распределение	44
5.8. Распределение Гаусса или нормальное распределение	45
5.9. Свойства нормального распределения	46
5.10. Линейная регрессия	47
5.11. Прогнозирование на основе методов линейной регрессии	50
Вопросы для самоконтроля	51
Список литературы	52
Лекция 6. Имитационное моделирование	53
6.1. Общие положения	53
6.2. Виды моделей	54
6.3. Построение моделей	54
6.4. Имитационные модели	55
6.5. Пример имитационного моделирования	55
6.6. Этапы имитационного моделирования	59
6.7. Выводы	61
Вопросы для самоконтроля	61
Список литературы	62
Лекция 7. Data Mining. Компьютерные технологии анализа данных	63
7.1. Современные технологии поиска и анализа данных	63
7.2. Технология Data Mining: понятие, назначение	63
7.3. Приложения технологии Data Mining	64
7.4. Новые сферы применения Data Mining	66
7.5. Классификация систем Data Mining	67
7.6. Развитие систем Data Mining	70
Вопросы для самоконтроля	71
Список литературы	71
Лекция 8. Информационная поддержка моделирования на основе документальных и инструментальных информационных систем	72
8.1. Электронные документы	72
8.2. Современные системы электронного документооборота	72
8.3. Информационно-поисковые системы	73
Вопросы для самоконтроля	74
Список литературы	74
Лекция 9. OLAP- механизмы современных систем поддержки принятия решений	76
9.1. Основные понятия	76
9.2. Компьютерные системы поддержки принятия решений	76
9.3. Характеристики систем поддержки принятия решений	77
9.4. Классификация СППР	77
Вопросы для самоконтроля	79
Список литературы	79
Лекция 10. Основы телекоммуникаций и сетевых технологий	80
10.1. Общие принципы построения компьютерных сетей	80
10.2. Локальные сети	80
10.3. Глобальные компьютерные сети	82
Вопросы для самоконтроля	83

Список литературы	83
Лекция 11. Виды угроз в информационно-коммуникационных системах	85
11.1. Угрозы объектам информационной деятельности	85
11.2. Компьютерные вирусы и антивирусные программы защиты	86
11.3. Классификация антивирусных программ	87
Вопросы для самоконтроля	87
Список литературы	88
Лекция 12. Установка, настройка и эксплуатация антивирусного программного обеспечения	89
12.1. Функции современного антивирусного программного обеспечения	89
12.2. Методы предотвращения вирусных атак	90
12.3. Источники вирусной опасности	91
12.4. Новые типы антивирусного программного обеспечения	91
12.5. Эвристический метод распознавания новых вирусов	92
12.6. Особенности применения антивирусных программ	93
Вопросы для самоконтроля	94
Список литературы	94
Библиографический список	96
Содержание	97