

В.С. МОИСЕЕВ

**ЛЕКЦИИ ПО МАТЕМАТИЧЕСКОМУ
МОДЕЛИРОВАНИЮ
В ПРИКЛАДНОЙ ИНФОРМАТИКЕ**



Казань 2018

Посвящается
100-летию победоносной
Советской (Красной) Армии
и орденоносного
Всесоюзного Ленинского
Коммунистического Союза
Молодежи (ВЛКСМ)

В.С. МОИСЕЕВ

ЛЕКЦИИ ПО МАТЕМАТИЧЕСКОМУ МОДЕЛИРОВАНИЮ В ПРИКЛАДНОЙ ИНФОРМАТИКЕ

Научно-учебное издание



Казань 2018

УДК 519.682.2:681.3.06
ББК 32.973
М 75

Редактор серии:

В.С. Моисеев – заслуженный деятель науки и техники Республики Татарстан, д-р техн. наук, профессор.

Моисеев В.С.

М 75 Лекции по математическому моделированию в прикладной информатике: научно-учебное издание. – Казань: Редакционно-издательский центр «Школа», 2018. 307 с.
(Серия «Современная прикладная математика и информатика»).

ISBN 978-5-906935-78-6

В книге приводится генезис информатики, ее связь с кибернетикой и развитием средств вычислительной техники. Дается современная классификация математических моделей, применяемых в науке, технике и экономике. Рассмотрен существующий и перспективный математический аппарат, который предлагается использовать в моделях информационных систем, технологий и в задачах защиты информации. Приведено значительное число математических моделей прикладной информатики.

Книга предназначена для магистрантов, аспирантов, докторантов, занимающихся исследованиями и разработками в области прикладной информатики. Материал лекций может быть полезен преподавателям ВУЗов при подготовке соответствующих лекций, лабораторных практикумов и выборе тем курсовых и выпускных квалификационных работ.

УДК 519.682.2:681.3.06
ББК 32.973

ISBN 978-5-906935-78-6

© Моисеев В.С., 2018
© РИЦ «Школа», 2018

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
Раздел I. ОСНОВНЫЕ ПОНЯТИЯ ИНФОРМАТИКИ И МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ПРОЦЕССОВ И СИСТЕМ	9
1. Возникновение основ информатики за рубежом и в СССР	9
2. Определение, история и направления современной информатики	20
3. Понятие и классификация информационных систем и информационных технологий.....	30
4. Определение и классификация математических моделей процессов и систем	40
Раздел II. ЭЛЕМЕНТЫ СОВРЕМЕННОЙ ТЕОРИИ ОТНОШЕНИЙ ДИСКРЕТНЫХ МНОЖЕСТВ.....	53
5. Основы классической теории отношений дискретных множеств.....	53
6. Теоретико-множественные операции над отношениями.....	63
7. Специальные операции над отношениями	70
8. Случайные бинарные отношения	77
9. Условные бинарные отношения	85
10. Динамические бинарные отношения.....	91
11. Стохастические бинарные отношения	102
Раздел III. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ИНФОРМАЦИОННЫХ СИСТЕМ	114
12. Бинарные модели информационных систем.....	114
13. Модели работы пользователей с банком данных информационной системы	126
14. Модель корпоративной информационной системы...	139
Раздел IV. МОДЕЛИ ПРИКЛАДНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.....	148
15. Модели прикладной информационной технологии ...	148

16. Модели управления перспективной прикладной информационной технологией интернет-рекламы.....	162
17. Модели распределенного банка данных прикладных информационных технологий	171
18. Модели комплексов программ прикладных информационных технологий.....	181
Раздел V. ЗАДАЧИ ОРГАНИЗАЦИИ РАЗРАБОТКИ И ЭКСПЛУАТАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ ..	202
19. Модели некоторых задач организации разработки информационных систем.....	202
20. Некоторые модели процессов эксплуатации информационных систем.....	217
Раздел VI. МОДЕЛИ СУЩЕСТВУЮЩИХ И ПЕРСПЕКТИВНЫХ ТЕХНОЛОГИЙ ЗАЩИТЫ ИНФОРМАЦИИ.....	242
21. Модель традиционной системы защиты информации.....	242
22. Теоретико-игровая вероятностная модель шифрования и дешифрования данных с применением алгоритмов DES и RSA.....	253
23. Дискретная модель активной защиты информационных систем.....	266
24. Марковская модель функционирования программных средств активной защиты информационных систем.....	279
25. Расчетная модель оценки информационной живучести АСУ специального назначения	290
ЗАКЛЮЧЕНИЕ	301
ИСПОЛЬЗОВАННАЯ И РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА.....	302

ВВЕДЕНИЕ

Двадцатый век характеризовался лавинообразным возрастанием объёмов информации и дефицитом времени на ее сбор и обработку. Следствием этого явилось появление в 40-50-х годах XX века специальных изделий приборостроения под названием электронно-вычислительные машины (ЭВМ).

В настоящее время информационные системы на базе ЭВМ используются в различных сферах деятельности общества. При этом широко обсуждаются вопросы об информационном обществе, информационной индустрии, информационных технологиях и т.д. Основу всех этих достижений составляет научная дисциплина – информатика, но при всех успехах практической информатики можно констатировать отсутствие к настоящему времени достаточно разработанных теоретических основ этой научно-технической дисциплины.

Основными причинами такого состояния являются:

1. Сложность и практическая неочевидность непосредственного использования методов «классической непрерывной математики» (дифференциальное и интегральное исчисления, дифференциальные уравнения и т.д.) из-за объективной дискретности информационных процессов.

2. Значительное многообразие используемых на практике видов информации, информационных систем и технологий, что не позволяет выбрать единую схему системы и технологии ее применения для их последующего математического описания.

Главной причиной, на наш взгляд, является отсутствие глубокого развития теоретической информатики, которая должна закладывать математические основы для формального анализа и синтеза информационных систем и информацион-

ных технологий. В качестве подтверждения этого тезиса можно указать на роль теоретической механики в развитии техники. Эта наука позволяет математически оперировать с такими понятиями, как материальная точка, твердое тело, система точек, скорость, ускорение, сила, моменты сил и др. Результаты этой дисциплины используется в таких технических дисциплинах, как динамика полета летательных аппаратов (ЛА), расчет на прочность конструкций и др., которые позволяют инженеру количественно обосновать принимаемые им проектные решения.

Целями данной работы являются:

1. Рассмотрение современного понятия информатики, как науки.
2. Обоснование необходимости применения математических моделей и методов, как базы для теоретических и прикладных исследований и разработок в области информатики.

Эти цели достигаются решением следующих основных вопросов, рассматриваемых в работе:

- история развития, современное содержание и основные понятия информатики;
- классификация современных математических моделей;
- основы теории отношений дискретных множеств;
- математические модели информационных систем и технологий;
- математические модели программного и информационного обеспечения;
- модели и методы защиты информации;

В основу книги положен курс лекций, прочитанных в Казанском государственном техническом университете им. А.Н. Туполева (КАИ) магистрантам направления «При-

кладная математика и информатика». Материал курса значительно расширен за счет включения в него результатов исследований и разработок автора и его учеников (к.т.н. Бутузова А.В., к.т.н. Тутубалин П.И., к.т.н. Козар А.Н., Гинатуллин И.А., к.т.н. Зиновьев П.А. и др.) в области применения математических моделей и методов в прикладной информатике.

Для удобства использования предлагаемый материал разбит на шесть разделов:

Раздел I. Основные понятия информатики и математического моделирования процессов и систем.

Раздел II. Элементы современной теории отношений дискретных множеств.

Раздел III. Математические модели информационных систем.

Раздел IV. Модели прикладных информационных технологий.

Раздел V. Задачи организации разработки и эксплуатации информационных систем.

Раздел VI. Модели существующих и перспективных технологий защиты информации.

Для углубленного изучения материала «Лекций», его применения и развития в книге приведен значительный список использованной и рекомендуемой литературы.

Отметим, что основной целью данной работы является привлечение внимания к применению математических моделей и методов в исследованиях и разработках в области прикладной информатики.

Книгу можно «скачать» и высказать замечания и предложения по ее содержанию на сайте: моисеев-бпла.рф, на кото-

ром можно ознакомиться с другими работами автора в области прикладной математики и информатики.

Автор благодарит к.т.н. Моисееву Л.Т. за большой труд по редактированию и подготовку рукописи книги к печати.

P.S.

Практически вся научная деятельность автора была связана с исследованиями и разработками в области применения прикладной математики и информатики в военно-технических и оборонно-промышленных задачах.

Автор вступил в члены ВЛКСМ в 1960 г. В 17 лет, работая на оборонном заводе, активно участвовал в работе боевой комсомольской дружины (БКД) Приволжского района г. Казани. Был бойцом студенческого строительного отряда Казанского авиационного института им. А.Н. Туполева. Избирался комсоргом группы. Работал внештатным инструктором Татарского обкома ВЛКСМ и руководил разработкой АСУ «Комсомол Татарии».

Активная жизненная позиция автора, воспитанная работой в комсомоле, отразилась в написании и опубликовании этой книги, а также в издании в течение 2004-2018 г.г. более 40 монографий, учебников и учебных пособий в задуманной и редактируемой им серии «Современная прикладная математика и информатика».

Раздел I. ОСНОВНЫЕ ПОНЯТИЯ ИНФОРМАТИКИ И МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ПРОЦЕССОВ И СИСТЕМ

Термин «информатика» и производные от него термины прочно вошли в повседневную жизнь нашего общества. Специалисты и неспециалисты к месту и не к месту говорят об «информационных технологиях (ИТ)», «ИТ-специалистах», «ИТ-индустрии», «информационном обществе», «информационном пространстве» и т.п.

Такая же ситуация с термином «кибернетика» имела место в 60-70 г.г. прошлого века. На наш взгляд, основной причиной ее забвения в настоящее время является тот факт, что кибернетика в своем развитии так и не стала четким научным направлением с общепризнанной теорией и практическими результатами.

Для того чтобы такая участь не постигла информатику, необходимо четко выявить ее истоки, сформулировать предмет и задачи информатики и, главное, очертить общепонятную математическую базу для выполнения исследований и разработок по решению ее теоретических и практических задач. Часть этих вопросов будет рассмотрена в лекциях данного раздела.

1. Возникновение основ информатики за рубежом и в СССР

Очевидно, чтобы глубоко понять суть любой науки или некоторого социального явления необходимо знать историю их появления и развития.

Прародительницей широко используемого в настоящее время термина «информатика» является кибернетика.

По определению кибернетика – наука об общих закономерностях процессов управления и передачи информации в системах различной природы (технических, биологических, социальных). Отцом кибернетики считается американский ученый Н. Винер, который в 1948 г. опубликовал книгу «Кибернетика или управление и связь в животном и машине».

Отметим, что слово «кибернетика» в переводе с древнегреческого означает «кормчий», т.е. человек, управляющий движением судна.

В основу этой науки Н. Винер положил проведенные во время второй мировой войны исследования по управлению огнем зенитной артиллерии по скоростным и маневрирующим воздушным целям. В этих процессах предусматривалось определение точки прицеливания орудия с упреждением, учитывающим движение цели и корректировка координат этой точки с учетом данных о ее маневрировании, полученных от следящей системы на базе радиолокационной станции.

Для реализации этих достаточно сложных задач были использованы электромеханические счетно-решающие устройства (СРУ) на основе релейных элементов. Такие устройства стали прародителями современной компьютерной техники.

К своим исследованиям Н. Винер привлек врачей и биологов, которые установили, что передача информации, обратная связь и управление присутствуют и в живых организмах.

С середины 50-х годов прошлого века поток зарубежных и отечественных статей и монографий по кибернетике резко увеличивается. На серьезное внимание к этой науке в СССР указывает создание в 1955 г. Научного совета по комплексной проблеме «Кибернетика» АН СССР. В последующем были созданы «Институт кибернетики» (г. Киев), «Институт техни-

ческой кибернетики» (г. Минск) и другие научные организации, решающие теоретические и практические задачи этого научного направления.

Однако следует отметить, что в нашей стране возникли и конъюнктурные направления этой науки (медицинская кибернетика, биокибернетика, военная кибернетика и др.), которые кроме общих фраз, заимствованных у Н. Винера и его зарубежных и отечественных последователей, не давали никаких практических результатов.

Усложнение задач управления, рассматриваемых кибернетикой, вызвало необходимость использования более мощных и быстродействующих СРУ, которая явилась объективной предпосылкой развития электронной вычислительной техники (ЭВТ). В рамках этого направления до 80-х годов XX века развивались 2 ветви ЭВТ – аналоговая и цифровая ВТ.

Первая реализовывалась в виде аналоговых вычислительных машин (АВМ) и комплексов (АВК), в которых все требуемые вычисления производились на их операционных усилителях путем изменения значений электрических напряжений. Аналоговая ВТ, имея масштаб вычислений реального времени, широко использовалась при полунатурных испытаниях аналоговых систем управления летательных аппаратов, при которых на АВМ (АВК) моделировались их полеты. Основным недостатком этого вида ЭВТ являлись значительные затраты времени на программирование задач, которое осуществлялось на наборном поле АВМ (АВК) совокупностью специальных проводов.

Поэтому цифровая ВТ вытеснила из повседневной практики аналоговую ВТ.

В 1942 г. в США была разработана первая в мире электронная цифровая вычислительная машина (ЭЦВМ) ЭНИАК стандартной структуры, предложенной Дж. Нейманом.

В 1948 г. постановлением Совета Министров СССР в рамках АН СССР был создан Институт точной механики и вычислительной техники (ИТМиВТ). В этом же году в нашей стране было зарегистрировано изобретение В.И. Рамеева и И.С. Брука первой отечественной ЭЦВМ. В 1950 г. под руководством С.А. Лебедева была создана первая в Европе ЭЦВМ МЭСМ (малая электронная счетная машина). Через 2 года появилась БЭСМ-1, которая в конце 50-х годов превратилась в самую мощную ЭВМ БЭСМ-6 на электронных лампах, имеющую быстродействие 1 млн. операций в секунду. В 1953 г. в СССР начато серийное производство ЭЦВМ «Стрела».

При этом в нашей стране были широко развернуты работы по кибернетике, развитию вычислительной математики и программированию для применения существующих и создаваемых ЭВМ при решении в СССР ракетной, космической и атомной проблем. В частности, в конце 50-х годов в Советском Союзе были разработаны алгоритмические языки Алгол-60 и Фортран, которые с различными модификациями просуществовали до середины 80-х годов прошлого века.

Бурное развитие в это время получило создание новых образцов вычислительной техники. Наряду с самой мощной ЭВМ БЭСМ-6 были разработаны следующие семейства машин 1-го, 2-го и 3-го поколений:

- большие ЭВМ – Урал-1, Урал-2, Урал-10, Урал-14;
- средние ЭВМ – М-2, М-20, М-220, М-222;
- средние ЭВМ – Минск-2, Минск-22, Минск-32;
- малые ЭВМ – Мир-1, Мир-2, Наири-С, Наири-К.

Отметим, что ЭВМ «Минск-32» имела первую в СССР операционную систему, которая называлась «лентосистема». В малогабаритных ЭВМ «Наири-С» и «Наири-К» кроме языка машинных команд для программирования задач использовался русскоязычный алгоритмический язык «АП» (алгоритмическое программирование). В отличие от отмеченных выше ЭВМ, у которых ввод данных и программ осуществлялся с помощью перфокарт и перфолент, взаимодействие пользователя-программиста с этими ЭВМ осуществлялось с использованием электрической пишущей машинки «Консул».

Кроме чисто вычислительных задач в СССР рассматривались задачи применения ЭВМ для непосредственного (прямого) управления различными видами оборудования. Для их решения в 1958 г. в составе АН СССР был создан Институт электронных управляющих машин (ИНЭУМ). В этом же году были разработаны управляющие вычислительные машины (УВМ) «Днепр» и МН-1 для задач АСУТП.

В 70-х годах прошлого века при развитии вычислительной техники по инициативе профессора МГУ М.Р. Шура-Бура в СССР был выдвинут лозунг: «Перегнать – не догоняя». Согласно этой идеологии в Советском Союзе свертывались все отечественные разработки ЭВМ, которые заменялись стандартными семействами машин с зарубежными прототипами:

1. Большие и средние ЭВМ семейства ЕС ЭВМ (Единое Семейство ЭВМ) с прототипами фирмы IBM.
2. Малые (управляющие) ЭВМ семейства СМ ЭВМ (Семейство Малых ЭВМ) с прототипами фирм PDP и HP.
3. Микро-ЭВМ семейства «Электроника» с прототипами различных фирм.

В их рамках были разработаны и запущены в производство ЭВМ ЕС1010, ЕС1020, ЕС1030, ЕС1040, ЕС1060, ЕС-1840, СМ-1, СМ-2, СМ-4, СМ-1420, «Электроника ДЗ-28», «Электроника-60», «Электроника-85». В последующем появились модификации ЭВМ семейства ЕС такие, как ЕС1022, ЕС1033, ЕС1045, ЕС1065 и др.

В рамках реализации этого лозунга был осуществлен переход на зарубежные алгоритмические языки Assembler, PL-1, COBOL и системное программное обеспечение ЭВМ-прототипов.

Отметим, что крупные научные организации АН и Минобороны СССР до конца 80-х годов решали свои задачи на ЭВМ БЭСМ-6 с использованием языков АЛГОЛ и ФОРТРАН.

В это время был выполнен ряд оригинальных разработок в области вычислительной техники. Был создан двухмашинный вычислительный комплекс ВК-33, в котором вторая ЭВМ ЕС1033 работала в «горячем резерве» при решении крупных задач. Для решения больших задач с очень высокой требуемой точностью был разработан четырехмашинный комплекс ВК-45, в котором задача одновременно решалась на трех ЭВМ ЕС1045 с периодической остановкой в заданных контрольных точках. В этих точках четвертая управляющая ЭВМ сравнивала полученные три результата и по принципу «голосования» выбирала из них наиболее точный результат с последующим продолжением решения задачи на трех ЭВМ комплекса.

Отметим, что отдельные составные части ЭВМ ЕС1033, разработанной СКБ Казанского завода ЭВМ под руководством В.Ф. Гусева закупались американскими фирмами.

Образцы ЕС1840 и «Электроника-85» являлись первыми серийными отечественными персональными ЭВМ. В 70-х годах была разработана первая в мире гибридная вычислительная система ГВС-100, включающая в себя взаимодействующие в вычислительном процессе аналоговый вычислительный комплекс и цифровую ЭВМ.

Специалистами кафедр ЭВМ и прикладной математики КАИ им. А.Н. Туполева был создан ГВК на базе АВК-31 и микро-ЭВМ «Электроника ДЗ-28», который показал значительно более высокое быстродействие гибридных вычислений по сравнению с цифровыми.

Развитие кибернетики и вычислительной техники в СССР было ориентировано на эффективное решение актуальных народнохозяйственных и оборонных задач.

Последние успешно решались в рамках следующих больших проблем:

- атомно-ядерной;
- ракетно-космической;
- радиоэлектронной и микроэлектронной.

В 1959 г. впервые в мире А.И. Китовым была выдвинута идея создания Единой Государственной Сети Вычислительных Центров (ЕГСВЦ) страны, объединяющей все ЭВМ Советского Союза. В 1964 г. был разработан проект ЕГСВЦ, включающий в себя 100 ВЦ, объединенных широкополосными каналами связи. К этим мощным ВЦ, которые выступали как центры коллективного пользования (ВЦКП) подключались 20 тыс. более мелких ВЦ. Отметим, что первая комплексная военная сеть ВЦ ARPANET в США, которая явилась прообразом интернета, начала функционировать только в 1969 г. Материалы проекта ЕГСВЦ с конкретизацией при-

кладного программного обеспечения были использованы в 1980 г. при разработке проекта Общегосударственной автоматизированной системы (ОГАС) планирования и управления народным хозяйством СССР, выполненного под руководством крупнейшего в СССР специалиста, заложившего основы безбумажной информатики, акад. В.М. Глушкова.

Отметим, что все предпосылки к созданию такой системы в нашей стране имелись. В частности в 1975 г. для Госплана СССР была разработана настольная (персональная) ЭВМ Искра-226, выпущенная в скором времени в количестве нескольких тысяч экземпляров. Эти ЭВМ вместе с большими ЭВМ БЭСМ-6 в 1958 г. составили комплекс технических средств автоматизированной системы плановых расчетов (АСПР), которая использовалась как подсистема ОГАС.

Для предприятий девяти ведущих министерств отечественной промышленности (Минавиапром, Минсудпром, Минрадиопром, Минэлектронпром и др.) была разработана типовая АСУП «Кунцево», совокупность которых предполагалось включить в состав ОГАС наряду с отраслевыми АСУ (ОАСУ) Министерств и ведомств СССР. Типовой проект последней был разработан и реализован в Минприборе СССР.

В качестве важной предпосылки внедрения ОГАС в народное хозяйство можно отметить значительное развитие в СССР и за рубежом в 60-70 г.г. такого направления прикладной математики как экономико-математические методы.

К сожалению, в силу ряда причин разработка и внедрение ОГАС были приостановлены. В настоящее время результатом этих работ является автоматическая междугородняя телефонная связь по кодам населенных пунктов РФ.

Укажем ряд оригинальных разработок, выполненных в нашей стране в «доперестроечный», «перестроечный» и «послеперестроечный» периоды:

1. В ВЦ АН СССР была создана система моделирования последствий для планеты ядерной войны. Полученными результатами активно заинтересовались в Ватикане и ООН.

2. Для ОКБ «Сухой» специалисты этого ВЦ разработали систему имитации воздушных боев разрабатываемых истребителей с существующими и перспективными истребителями вероятного противника.

Отметим, что остальные авиационные ОКБ в это время занимались автоматизацией проектных расчетов и оформления чертежей. В основе разработки этой системы лежит высказывание П.О. Сухого: «У меня в ОКБ достаточно инженеров, которые могут провести все необходимые расчеты на калькуляторах. Дайте мне систему, на которой я могу «разыгрывать» возможные воздушные бои будущей машины, анализировать результаты и корректировать ее параметры и характеристики при «завязке» проекта».

Как следствие такого подхода к применению ЭВМ сейчас истребители семейства «Су» являются лучшими истребителями в мире.

3. В 70-80 г.г. под руководством акад. В.С. Семенихина на базе комплексов ВК-33 и специальной периферийной техники была разработана АСУ Советской Армии (АСУ СА). Эта система с помощью защищенной помехоустойчивой и быстродействующей системы связи, развернутой на территории стран Варшавского Договора, объединила штабы военных округов, родов войск и крупных соединений, в том числе систему управления ПРО г. Москвы.

4. В ИТМиВТ в 70-90 г.г. было создано семейство много-процессорных супер-ЭВМ «Эльбрус» (10-16 процессоров; 12,5 млн.оп./с; 72 разряда машинного слова; 144 Мб оперативной памяти) и мощные ОС «Диспак» и «Дубна». При этом одна ЭВМ «Эльбрус-2» заменяла работу трех ЭВМ БЭСМ-6. Отметим, что американская супер-ЭВМ «Cray» имела в то время быстродействие 5 млн. оп./с.

5. Серийно выпускалась персональная ЭВМ ЕС1845 с защитой ее системного блока от перехвата излучений и программируемой клавишей, позволяющей мгновенно стирать содержимое ее памяти.

6. Специалистами ОКБ «Корунд» было разработано семейство средств вычислительной техники специального назначения БАГЕТ, включающего в себя персональные, бортовые одноплатные ЭВМ и серверы, а также полный набор периферийных устройств для его эксплуатации в стационарных и полевых условиях.

7. Были разработаны защищенные мобильная операционная система МСВС и система управления базами данных ЛИНТЕР.

Во времена «перестройки», под разговоры о том, что в СССР нет, не было и не будет хорошей вычислительной техники, в нашу страну хлынул поток крайне слабых зарубежных ПЭВМ типа IBM PC, ICL и др.

Последствия этих «сказок» ощущаются и в настоящее время. В лучшем случае в современной России производство аппаратных средств информатики ведется по «отверточной технологии» из иностранных (китайских) комплектующих.

Широко распространенный в настоящее время ИНТЕРНЕТ ориентирован на общество потребителей, но не созида-

телей и является не более чем средством оперативной почтовой и телефонной связи, большой информационно-справочной системой, магазинами, кинотеатром, теле- и радиостанциями, концертным залом и т.п. При этом процессы информатизации различных сторон практической деятельности общества в нашей стране (корпоративные системы, ERP-системы, CAD/CAM/CAE-системы, порталы государственных и других видов услуг, «электронное правительство», «облачные вычисления» и др.) проводятся бессистемно с бездумным использованием зарубежных платформ и оболочек.

Эти работы ведутся при отсутствии глубокой научной базы в лице информатики, которую ранее составляла такая научная дисциплина как кибернетика. Отметим такой парадокс, что последняя практически не упоминается в современной научно-технической литературе, но при этом широко используются термины: «киберпреступность», «кибербезопасность», «кибератаки», «киберкомандование», «кибервойска» и т.п.

Вместе с тем в Российской Федерации в конце 1990-х и в начале 2000-х годов был создан ряд научно-исследовательских организаций, в задачи которых входит развитие современной информатики. В частности к ним относился Институт проблем информатики АН Республики Татарстан (ИПИ АН РТ), с которым автора связывало многолетнее сотрудничество.

В заключение лекции отметим, что в СССР и РФ до 2000-х годов были созданы достаточно мощные отечественные аппаратные и программные средства информатики. К сожалению, в этот период не были разработаны общепринятые теоретические основы информатики, которые при соответствующем развитии можно было положить в основу современных прикладных исследований и разработок информационных систем и технологий.

2. Определение, история и направления современной информатики

Информатика – наука, изучающая все аспекты получения, хранения, преобразования, передачи и использования информации. Как следует из Лекции №1, информатика начала активно развиваться с середины XX века, зародившись в недрах другой обобщенной науки – кибернетики (Рис. 1).

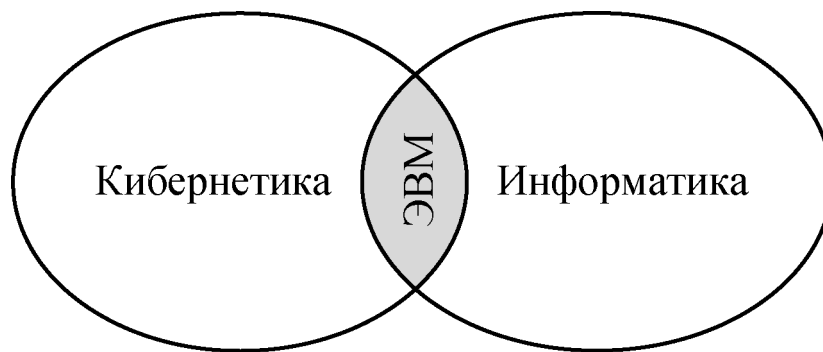


Рис. 1

Как следует из этой лекции, в кибернетике, как во всеобщей науке об управлении, активно использовались ЭВМ. С усложнением задач управления, которые были связаны с обработкой больших объёмов информации, от неё отпочковалась специальная дисциплина - информатика.

Информатика имеет достаточно древние корни. Условно можно выделить следующие поколения информатики:

«**Информатика 0**», которая связана со сбором, хранением, преобразованием, передачей информации в летописях, на папирусах и т.п.

«**Информатика 1**» – (сер. 50-х-сер. 60-х годов XX века).

Другое ее название – «механическая информатика». Процессы получения, хранения информации были связаны с активным книгопечатанием. Передача информации осуществлялась почтой, а ее поиск и обработка – с помощью механических устройств (штырьевые устройства отбора карт, арифмометры).

«**Информатика 2**» – (сер. 60-х-конец 70-х годов XX века), которая позиционируется как «электромеханическая информатика», применяющая на практике счетные машины, телефон, телеграф и т.п.

«**Информатика 3**» – (начало 80-х годов XX века).

Этот этап развития получил название «электронная информатика», связанная с появлением различных классов ЭВМ с их периферийными устройствами (АЦПУ, дисплей, графопостроитель и т.п.).

Термин «информатика» появился в «Информатике 1» как составная часть библиотечного дела, но с появлением «Информатики 3», выделился в самостоятельную науку, которая получила название *informatique*, что, в переводе с французского, означает «информационная автоматика».

К настоящему времени в составе информатики предлагается выделить несколько основных направлений, которые представлены на Рис. 2.



Рис. 2

Философские вопросы информатики исследуют процессы возникновения и развития понятий «информация» и «информатика», а также их роль в современной и будущей жизни общества.

Как сказал кто-то из древних мыслителей, основная задача философии – дать четкие представления о туманных мыслях обычных людей. А как показала современная практика, таких мыслей о возможностях информатики существует великое множество.

Теоретическая информатика – это математическая дисциплина, которая использует формальные методы для построения и исследования моделей обработки, передачи и использования информации.

Прикладная информатика – это область техники, связанная с решением конкретных задач информатизации всех сфер деятельности общества.

Остановимся более подробно на этом понятии, которое не получило к настоящему времени широкого распространения.

Современный этап развития общества характеризуется у нас в стране и за рубежом активным переходом от компьютеризации к информатизации практически всех сфер его деятельности. При этом главным его препятствием, на наш взгляд, является отсутствие глубоко проработанных теоретических основ информатизации деятельности человека, организации, региона, государства, общества в целом и, в частности, теории оптимального проектирования, эксплуатации и развития больших и сложных информационных систем, в которые внедряются информационные технологии.

Основу такой теории и технологий должны составить современные модели, методы и средства прикладной математики и информатики. При этом рассмотрение вопросов автоматизации формирования и оптимизации всевозможных решений должно ориентироваться на широкое применение математических моделей, методов и алгоритмов, реализуемых в соста-

ве соответствующих прикладных информационных технологий.

К настоящему времени нельзя говорить об информатике как о завершенной области научного знания с четко определенными предметом, целями, задачами и методами исследований. При этом практически отсутствуют научно обоснованные рекомендации по организации эффективной разработки, внедрения, эксплуатации и развития создаваемых информационных систем. Выход из этого положения видится нам в опережающем развитии такой ее составляющей, как «прикладная информатика».

Главной целью прикладной информатики является создание инженерных методик разработки современных информационных систем и технологий различного назначения. В этих методиках должны найти широкое применение современные и перспективные модели и методы прикладной математики, языки программирования, инструментальные средства и технологии разработки защищенных программ и баз данных, операционные системы и среды, системы управления базами и банками данных, аппаратно-программные средства хранения, обработки и передачи информации.

Активное развитие прикладной информатики позволит обобщать получаемые результаты в рамках соответствующих теорий, а это будет являться стимулом к развитию, как теоретической информатики, так и прикладной математики. Полученные при этом модели и методы будут использоваться в соответствующих методиках создания, эксплуатации и развития информационных систем и технологий их функционирования.

Рассмотрим краткую характеристику этих направлений.

Информация, которая бывает символьной и графической, в основном тяготеет к дискретным представлениям. Множество любых информационных сообщений также обычно является дискретным множеством. Отсюда следует, что теоретической базой информатики в основном является дискретная математика, но для некоторых задач анализа и синтеза информационных систем используются и методы «непрерывной математики» (например, теория массового обслуживания).

К настоящему времени к теоретическим основам информатики относят следующие научные дисциплины:

1. Дисциплины, основанные на математической логике (например, теория алгоритмов, теория вычислительных процессов, теория автоматов, теория параллельных вычислений).

2. Вычислительная математика и вычислительная геометрия, которые объединяют в себе методы численного решения на ЭВМ сложных задач и формального (математического) представления сложной графической информации.

3. Теория информации и кодирования, изучающая информацию в виде абстрактного объекта, а также методы представления информации для ее эффективной передачи и хранения.

4. Моделирование процессов сбора, хранения, обработки, передачи информации в реальных системах (например, сети Петри и т.п.).

5. Теория принятия решений, которая разрабатывает методы использования информации для выбора нужного решения из множества альтернативных возможностей. В ней широко используются методы оптимизации решений по совокупности заданных критериев. При этом для задач информа-

тики основную роль должны играть модели и методы дискретного программирования.

Перспективными задачами теоретической информатики являются:

1. Теория информатизации общества.
2. Теория Интернета с автоматическим и диалоговым уточнением запросов пользователя с использованием методов искусственного интеллекта.
3. Теория больших и сверхбольших баз данных с извлечением из них знаний и сверхбыстрым извлечением требуемых данных.
4. Теория автоматического программирования для ЭВМ с речевым и рукописным вводом заданий и исходных данных.
5. Теория описания, извлечения и обобщения знаний для систем искусственного интеллекта.

Прикладная информатика в настоящее время включает в себя следующие направления практической деятельности:

- анализ и программирование потоков разнообразной информации, циркулирующей в обществе. Данная задача является важной, но сейчас практически не исследуется.
- исследования и разработки способов представления, хранения и передачи информации (работы по базам данных, распределенным банкам данных, телекоммуникационным системам и т.п.).
- создание эффективных программных и аппаратных средств работы пользователей и разработчиков информационных систем.
- разработка сетей хранения, обработки и передачи информации.

- Разработка эффективных методов и средств обеспечения информационной безопасности.

Сформулируем основные перспективные задачи этого направления информатики:

1. Разработка «умных» операционных систем и систем управления базами данных.

2. Разработка платформ и оболочек для различных этапов и задач информатизации общества (государства).

3. Разработка программно-аппаратных средств систем моделирования социально-экономических процессов в обществе (государстве), в частности последствий принимаемых законов, указов и распоряжений.

4. Разработка типовых АСНИ, САПР, АСУП и АСУТП с оптимизацией исследовательских, проектных, управленческих и управляющих решений.

5. Разработка аппаратных и программных средств реализации аддитивных технологий производства, в частности изготовления из различных металлов деталей на 3D-принтерах.

6. Разработка программного обеспечения систем искусственного интеллекта для применения в составе АСНИ, САПР, АСУП, АСУТП и других видах информационных систем.

7. Разработка методов организации процессов создания программных и аппаратных средств сбора, хранения, обработки и передачи информации.

В заключительной части лекции остановимся на проблеме информатизации общества (государства).

Развитие процессов информатизации и инфокоммуникационных технологий (ИКТ) является приоритетной задачей любого цивилизованного общества. Повсеместное внедрение ИКТ представляется сегодня одним из важнейших факторов

усовершенствования сферы государственного управления, экономической и социальной сфер, повышения качества жизни населения страны в целом и отдельных ее регионов в частности.

Единая государственная политика в области информатизации должна строиться на базе типовых «сквозных» проектов, охватывающих как федеральный, так и региональный уровни управления. При этом приоритет должен отдаваться именно региональным аспектам развития ИКТ, что связано с необходимостью сосредоточения и концентрации имеющихся ресурсов на задачах создания региональной инфокоммуникационной среды, обеспечивающей любому российскому гражданину реальную возможность получения необходимых информационных ресурсов и услуг в любой точке России. В силу ряда причин путь информатизации России через информатизацию регионов является наиболее рациональным и перспективным.

Вместе с тем, анализ результатов в области информатизации, полученных у нас в стране, показывает, что эти процессы на региональном уровне по-прежнему ведутся стихийно, при отсутствии более или менее разработанной теоретической базы региональной информатизации. Последнее может привести к ошибкам и просчетам, которые имели место в 1970-80 гг. при таком же стихийном внедрении АСУ, САПР и АСУТП на базе эмпирических подходов. Декларируемое сейчас использование информатики как полноценной научной базы информатизации на данном этапе представляется не вполне обоснованным из-за известной аморфности предмета и недостаточной глубины проработки теоретических методов и подходов в рамках этого важного направления современной науки.

Одной из важнейших задач теории информатизации является разработка стратегии этого сложного многоступенчатого процесса, который должен охватить все уровни – от индивидуальной сферы деятельности человека до информационного обеспечения жизнедеятельности общества в целом. При этом общепринятым является утверждение, что переход к следующему, более высокому уровню информатизации должен происходить на основе «интеграции» всех компонентов предшествующего уровня в определенную систему (стратегия «снизу-вверх»). Для ускорения процесса интеграции и, в конечном итоге, для обеспечения оперативной и согласованной информатизации необходимо дополнить отмеченную выше стратегию процессом «декомпозиции» или «дифференциации», т.е. стратегией «сверху-вниз». Это подразумевает «нисходящую» структуризацию целей и ограничений, а также предварительную разработку «оболочки» системы более высокого уровня, заполнение имеющихся в ней «слотов» элементами (подсистемами) нижележащего уровня и настройку параметров их функционирования, исходя из необходимых условий достижения целей информатизации рассматриваемого иерархического уровня.

Как следует из вышеизложенного, общая стратегия информатизации РФ должна ориентироваться на информатизацию ее субъектов (регионов) с последующей интеграцией «электронных регионов» в «электронную Россию». Процесс информатизации целесообразно осуществлять на семи основных уровнях, представленных на Рис. 3, с помощью взаимосвязанных и взаимодополняющих процессов «интеграции» и «декомпозиции» («дифференциации») разрабатываемых информационных систем и ИКТ.



Рис. 3

Здесь «интеграция» подразумевает объединение отдельных систем нижнего уровня в единую интегрированную информационную систему более высокого иерархического уровня. «Дифференциация» имеет целью упорядочить информатизацию объектов нижнего уровня путем структурно-функциональной декомпозиции разрабатываемой системы более высокого уровня. На практике процесс «дифференциации» предполагается реализовать путем разработки типовых фрагментов и/или компонентов интегрированной информационной системы рассматриваемого уровня и их настройки на определенные задачи информатизации с учетом функций и задач объектов (систем и подсистем) низлежащего уровня.

Как следует из Рис. 3, региональная информатизация осуществляется на шести уровнях предлагаемой иерархии стратифицированных описаний объектов. В работе [Зиновьев П.А., Моисеев В.С., Фазылзянов Ф.М.] показано, что совместное применение «интеграции» и «декомпозиции» («дифференциации») информационных систем на этих уровнях позволяет достичь сокращения затрат времени и экономии финансовых средств при информатизации региона.

Теоретические вопросы информатизации службы скорой медицинской помощи (3-й уровень деятельности) изложены в книге [Бутузова А.В., Моисеев В.С., Тутубалин П.И.].

3. Понятие и классификация информационных систем и информационных технологий

Информационная система — это взаимосвязанная совокупность средств, методов и персонала, используемых для сбора, хранения, обработки и передачи информации в интересах достижения поставленных перед ней целей.

Любая информационная система должна удовлетворять следующим свойствам:

1. Информационная система должна быть проанализирована, построена и управляема на основе общих принципов.

2. Любая информационная система является динамической и развивающейся системой.

3. При построении любой информационной системы должен быть использован системный подход.

4. Выходными воздействиями информационной системы являются данные, используемые в общем случае для принятия решений.

5. Любая информационная система должна рассматриваться как интерактивная (человеко-машинная) система обработки информации.

Типовой состав автоматизированной информационной системы (АИС) представлен на Рис. 4.

Современные автоматизированные информационные системы по функциональному признаку подразделяются на две большие группы:

1. По сферам применения:

а) автоматизированные информационные системы организационного управления (бывшие автоматизированные системы управления (АСУ) производством);

б) автоматизированные информационные системы управления технологическими процессами (бывшие АСУТП);

в) CAD/CAM/CAE (ранее называемые системами автоматизированного проектирования (САПР));

г) интегрированные автоматизированные информационные системы, которые, по сути, являются объединением трех предыдущих видов систем. Ранее такие системы назывались

интегрированными АСУ (ИАСУ). Последнее время данный вид систем получил название корпоративных, (территориально разнесенных информационных) систем (КИС).



Рис. 4

2. По характеру используемой информации:

- а) информационно-поисковые (справочные) системы;
- б) информационно-решающие (управляющие, советующие) системы.

Отметим, что в настоящее время отсутствуют четкое определение и классификация такого широко применяемого на практике термина как «информационная технология». Поэтому детально рассмотрим понятия, этапы развития, классификацию информационных технологий (ИТ).

Процесс — это определенная совокупность действий, направленная на достижение поставленной цели.

Термин «технология» произошёл от греческого *techné* (искусство, мастерство, умение).

Понятие «технология» появилось и широко используется в сфере материального производства, где технология определяется как совокупность средств и методов обработки, изготовления, изменения состояния, свойств, формы сырья и материалов.

Схематически технология материального производства представлена на Рис. 5.

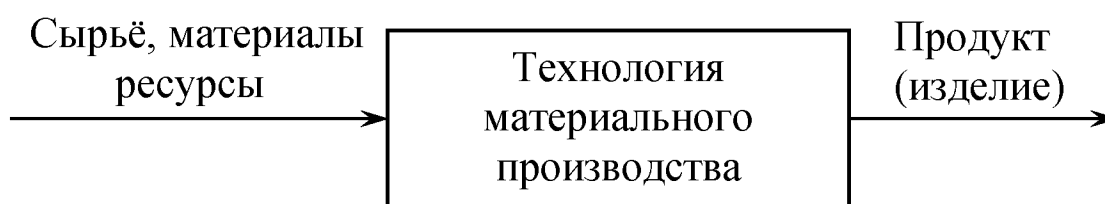


Рис. 5

Целью любой «материальной технологии» является получение материального продукта, удовлетворяющего определенные потребности человека и общества.

Как известно, информация является одним из важнейших ресурсов общества наряду с такими ресурсами, как нефть, газ, полезные ископаемые и др. Это означает, что процесс переработки информации также должен удовлетворять определённым технологиям. Схематично информационные технологии представляются Рис. 6.



Рис. 6

Информационная технология – процесс, использующий совокупность методов и средств сбора, хранения, обработки и

передачи данных для получения информации нового качества о состоянии объекта, процесса или явления (информационного продукта). Целью информационной технологии является производство информации для её анализа человеком и принятия решений с целью выполнения определённых действий.

Развитие информационных технологий включает в себя следующие этапы:

Этап 1. (до середины 19 века) – «Ручные информационные технологии» (перо, чернильница, книги, механические счетные приборы). Передача информации осуществлялась с помощью почты в виде писем, пакетов и т.д.

Этап 2. (с конца 19 века) – «Механическая информационная технология» (арифмометр, пишущая машинка, телефон, телеграф).

Этап 3. (до 50-х годов XX века) – «Электрическая информационная технология» (электрические калькуляторы, электрические пишущие машинки, копировальная техника и т.п.).

Этап 4. (60-80-ые годы) – «Электронные информационные технологии» (большие, средние и малые ЭВМ, сети связи, программные комплексы, базы данных и т.п.). На данном этапе было начато широкое применение математических методов для проведения всевозможных расчетов, оптимизации решений и т.п.

Этап 5. (середина 80-х годов и далее) – «Новая информационная технология» связанная с развитием персональных ЭВМ, средств телекоммуникаций и т. п. Основным результатом данного этапа является персонализация процессов сбора, хранения и обработки информации.

Характерными свойствами современных информационных технологий являются:

- интерактивность характера работы пользователя системы;
- интегрированность технических и программных средств системы;
- гибкость процесса изменения, как данных, так и постановок решаемых задач.

Реализация технологических процессов материального производства осуществляется с помощью станков, инструментов, специального оборудования, рабочей силы и производственных инструкций.

Любая современная информационная технология реализуется в среде определённой информационной системы. Следуя составу автоматизированной информационной системы, представленной на Рис. 4, можно выделить средства реализации информационной технологии, приведенные на Рис. 7.

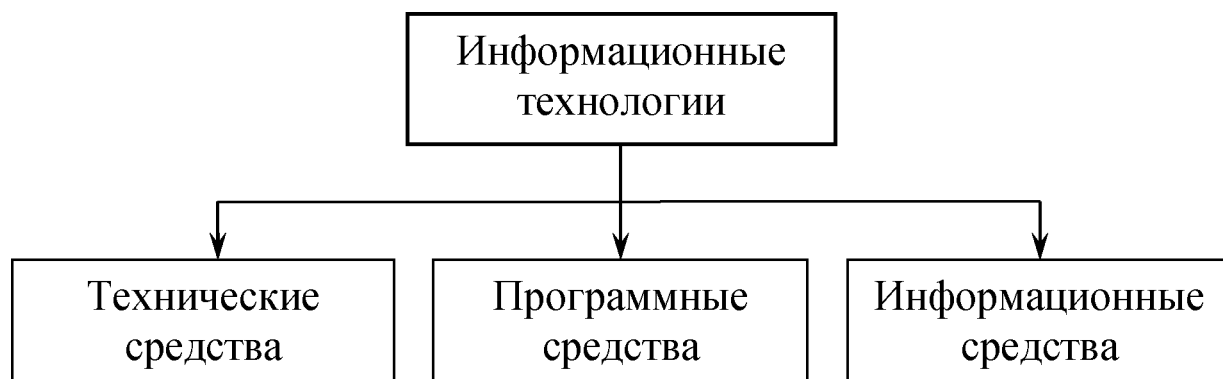


Рис. 7

Математические методы, модели, и алгоритмы переработки информации составляют математическое обеспечение соответствующей информационной технологии.

В общем случае любая информационная технология должна сопровождаться инструкциями пользователю и обслуживающему персоналу.

На Рис. 8 представлен типовой состав информационной технологии (ИТ):

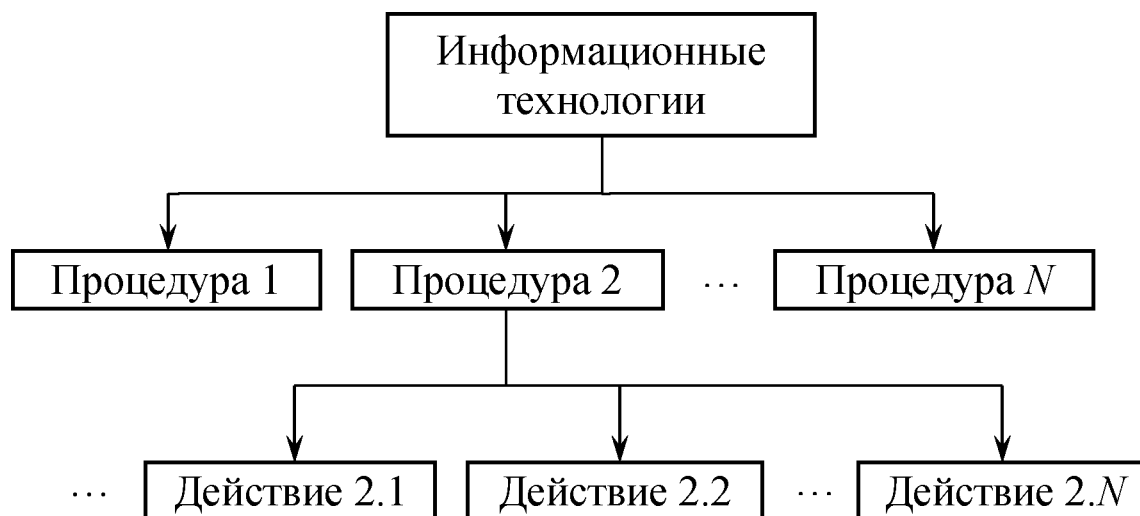


Рис. 8

Процедурами ИТ могут быть отдельные этапы, задачи, шаги процесса сбора, хранения, обработки и передачи информации. Для выполнения любой процедуры могут быть использованы одни и те же или разные средства, представленные на Рис. 7.

Действия выполняются пользователем информационной технологии по определённым инструкциям. При решении соответствующих задач в рамках конкретной технологии входящие в нее процедуры могут выполняться в различных последовательностях.

Рассмотрим классификацию существующих информационных технологий.

В наиболее широком смысле под ИТ понимается «совокупность способов реализации информационных процессов в различных областях человеческой деятельности при производстве информационных продуктов».

Для более детального анализа введем в рассмотрение специальную классификацию ИТ, представленную на рис. 9.



Рис. 9

Частные ИТ представляют собой упорядоченную последовательность элементарных технологических операций (процедур), осуществляемых в ходе выполнения информационных работ и носящих локальный характер. Примерами таких технологий являются ИТ сбора данных о ходе производства на предприятии, передачи информации между субъектами и объектами ИС, вывода данных на печать и т.д.

Совмещенные ИТ включают в себя отдельные взаимодействия между собой частные ИТ, имеющие целью выполнение определенных функциональных задач в ИС, связанных с поддержкой тех или иных решений. Например, реализация функциональной задачи «ведение базы данных (БД)» осуществляется с помощью следующих частных ИТ: «сбор (ввод) данных», «контроль достоверности данных», «запись данных в файлы БД», «контроль правильности записи в файлы», «модификация записей БД» и некоторых других. При работе с удаленными базами данных в их состав добавляется частная ИТ «передача данных».

Прикладные ИТ – совокупность взаимодействующих частных и совмещенных ИТ, необходимых и достаточных для принятия локальными и удаленными пользователями конкретных решений, связанных со сферами их деятельности. Примерами таких ИТ являются технологии планирования показателей производства, формирования социально-экономи-

ческих показателей региона, выбора проектных решений при создании новых сложных изделий и т.д.

Для совмещенных ИТ и особенно для прикладных ИТ, функционирующих в инфраструктурах региона и государства, существенную роль играет понятие информационно-коммуникационных технологий (ИКТ).

Под ИКТ будем понимать совокупность информационных технологий, предназначенных для сбора, обработки и хранения как традиционных данных, так и мультимедийной информации, а также их передачи на значительные расстояния. Роль ИКТ возрастает при объединении частных ИТ в совмещенные ИТ и совмещенных ИТ в прикладные ИТ, поскольку количество взаимодействующих субъектов и объектов инфраструктур увеличивается, и это, в свою очередь, неминуемо ведет к усложнению методов взаимодействия между ними.

Информационные технологии также являются развивающимися объектами, этапы и направления развития которых представлены на Рис. 10.

Модификация ИТ осуществляется путем исключения или добавления в существующие технологии операций (процедур), совершенствования применяемых алгоритмов и методик их выполнения, замены программных и аппаратных средств на более эффективные образцы и т.п. Разработка новых ИТ связывается с накоплением знаний и появлением новых задач. Отметим, что развитие ИТ должно обязательно подчиняться требованиям их устойчивого развития.

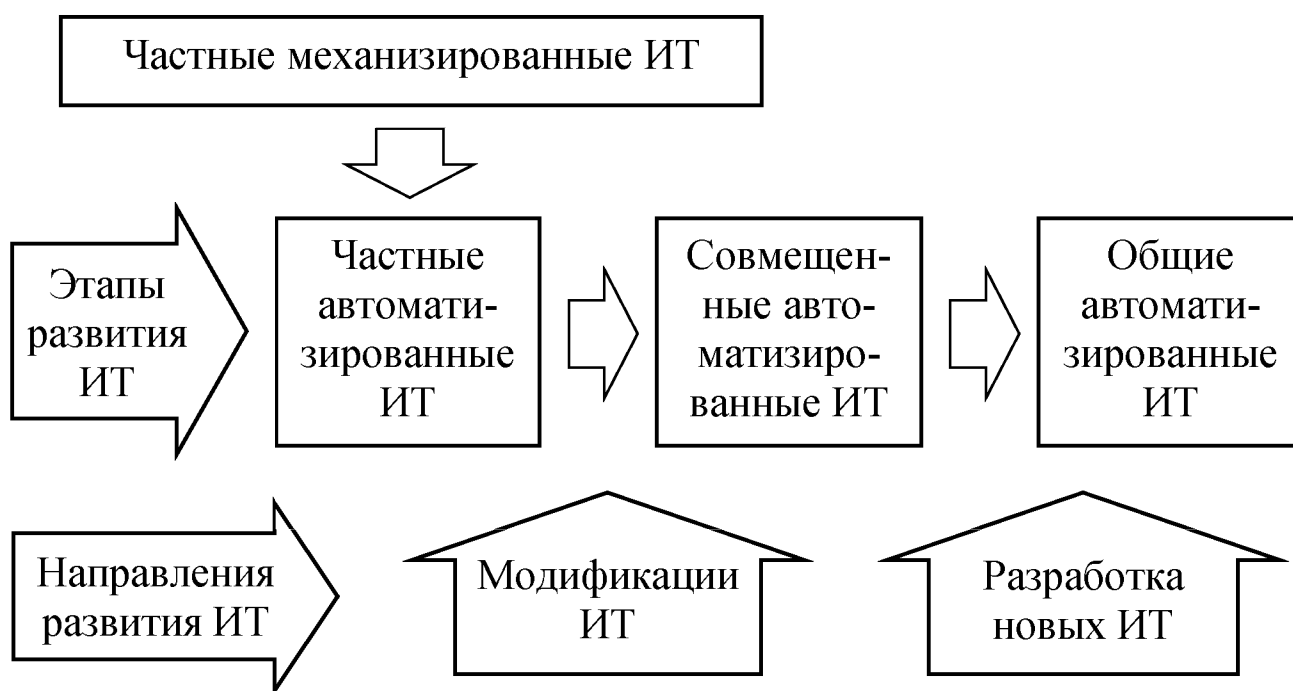


Рис. 10

В связи с тем, что все технологии функционируют в рамках определенной инфраструктуры, можно говорить о системах ИТ. Основными системообразующими факторами для совокупности частных, совмещенных и прикладных автоматизированных ИТ являются:

- 1) связи между ИТ типа «вход-вход», когда несколько технологий используют общие исходные данные;
- 2) связи между ИТ типа «выход-вход», когда результирующие данные некоторой ИТ используются другими технологиями;
- 3) связи типа «выход-выход», при которых несколько ИТ выдают свои результирующие данные в общую базу (хранилище) данных или определенной группе пользователей;
- 4) многосторонние связи между ИТ по совместно используемым каналам связи, средствам обработки данных, базам и банкам данных, пользователям и т.п.

Все эти факторы необходимо учитывать как при проектировании ИТ, так и при принятии соответствующих решений по развитию информационных систем.

В заключение лекции выделим следующие типы ИТ:

1. Технология обработки данных, для которой известны все необходимые данные и четко определены механизмы их обработки.

2. Информационная технология делопроизводства.

3. Информационная технология управления.

4. Информационная технология поддержки принятия решений, основанная на базах данных и базах моделей задач принятия решения.

4. Определение и классификация математических моделей процессов и систем

Математическое моделирование в науке, технике и производстве позволяет до создания реальных систем (объектов, процессов) или возникновения реальной ситуации в некотором процессе рассмотреть возможные режимы их работы, выбрать требуемые управляющие воздействия и составить объективный прогноз будущих состояний системы, объекта, процесса.

Вычислительные эксперименты, проводимые на основе математических моделей, помогают увидеть за частным общее, развить универсальные методы управления объектами и процессами различной природы, познать их общие закономерности функционирования.

Кроме этого, математическое моделирование является центральным звеном рассмотренных в Лекции №3 автоматизированных систем проектирования, управления и обработки данных.

Основная научная задача математического моделирования – выделение законов в природе и обществе и запись их на языке математики.

Существует множество определений математической модели. Приведем одно из них, на наш взгляд наиболее полно отражающее цели моделирования.

Математической моделью некоторого объекта, процесса или явления будем называть запись его свойств на формальном языке с целью получения нового знания (свойств) о них путем применения формальных (математических) методов.

Модель – это инструмент, ориентированный в первую очередь на исследование поведения и свойств конкретного объекта в целях управления этим объектом или предсказания его свойств.

Альтернативой формальному (математическому) подходу является экспериментальный подход. К его недостаткам можно отнести:

- высокую стоимость подготовки и проведения детальных экспериментов;
- получение частного знания (знания о конкретном объекте исследования, а не о классе объектов).

Пусть, например, требуется определить воздействие x на некоторый процесс или объект, при котором его результирующая характеристика y имеет максимально возможное значение (Рис. 11).

На Рис. 11, *а* показан эмпирический (экспериментальный) подход к решению поставленной задачи, который состоит в экспериментальном определении значения параметра y для нескольких значений входного воздействия x_1, x_2, \dots, x_8 . Среди них находится наибольшее, и оно принимается за максимум.

Как видим из этого рисунка, возможно несколько значений воздействия x (x_4 или x_5), при которых y имеет наибольшее значение, равное y_4 , но ни одно из них не является настоящим максимумом, который, возможно, лежит между ними.

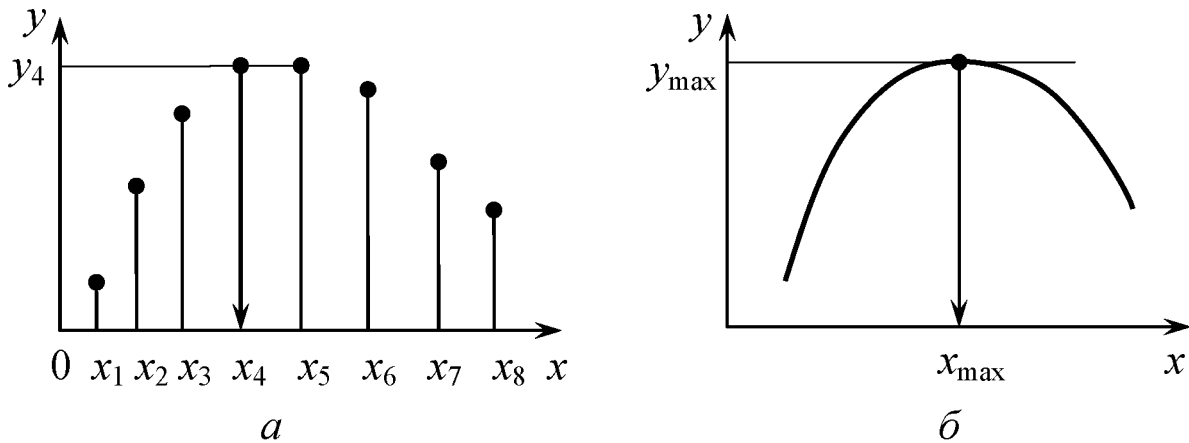


Рис. 11

Математический подход (Рис. 11, б) предполагает наличие математической модели процесса вида:

$$y = f(x).$$

Взяв производную $\frac{df}{dx}$ и приравняв ее к нулю, получим уравнение, решением которого является точное значение x_{\max} , доставляющее максимум функции y .

Схема применения математической модели при решении реальных задач имеет вид, показанный на Рис. 12.

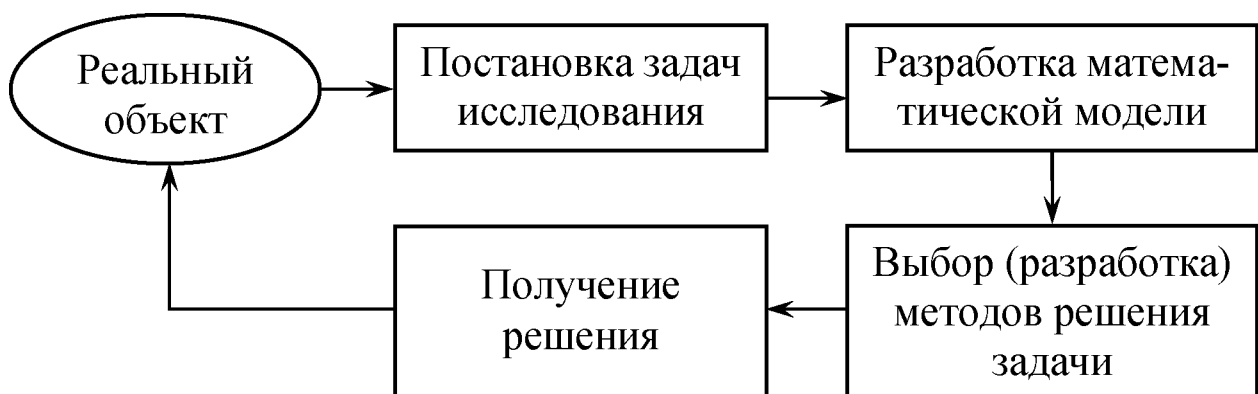


Рис. 12

Из этого рисунка следует, что процесс использования математических моделей реальных объектов и процессов является итерационным процессом, когда последовательно уточняются (дорабатываются) как математическая модель, так и методы решения стоящих задач.

Как показала практика разработки и применения разнообразных математических моделей, к ним предъявляются следующие основные требования:

- простота и наглядность модели;
- достаточная сложность модели;
- достаточная точность (адекватность) модели.

Простота и наглядность модели необходима для ее верификации исследователем и получения аналитического или численного решения с минимальными затратами времени.

Достаточная сложность модели обеспечивает возможность получения при ее реализации нетривиальных практически значимых результатов и выводов.

Важнейшей характеристикой построенных моделей является *точность*, обеспечивающая их адекватность реальности. При этом важно иметь в виду, что все модели представляют собой приближенное описание реальных объектов (процессов) и поэтому принципиально неточны. Интегральная оценка точности модели может быть получена путем сравнения результатов моделирования и экспериментальных данных для конкретных объектов и процессов.

Для оценки значимости совпадения или несовпадения модельных и экспериментальных результатов широко используются методы математической статистики. Вместе с тем не следует переоценивать результаты такой проверки. Хорошее совпадение модельных и экспериментальных данных, вообще

говоря, не доказывает точности модели, а лишь подтверждают ее функциональную пригодность для моделирования. Всегда может быть предложена другая модель, обеспечивающая лучшее совпадение с экспериментом, но не лучшее описание моделируемого объекта или процесса.

Выполнение приведенных выше противоречивых требований к математическим моделям достигается путем выполнения следующих общепринятых практических рекомендаций:

1. Включение в модель только основных параметров и характеристик рассматриваемого явления, процесса, объекта, системы.

2. Использование при построении модели формальных закономерностей, математически описывающих взаимосвязи параметров и характеристик моделирования, выбранных на этапе 1 (Рис. 12).

При недостаточной точности полученной модели на последующих этапах моделирования проводится корректировка состава параметров и характеристик объекта моделирования и (или) описывающих его закономерностей.

В настоящее время существует несколько видов классификации математических моделей, которые являются достаточно условными. Поэтому, говоря о классификации моделей, способах и схемах моделирования, необходимо очертить конкретную область исследования, чтобы отразить особенности моделируемых процессов. На Рис. 13 представлена схема, которая по нашему мнению, отражает классификацию моделей достаточно большой совокупности процессов различной природы.

По характеру решаемых реальных задач модели можно разбить на 2 класса:

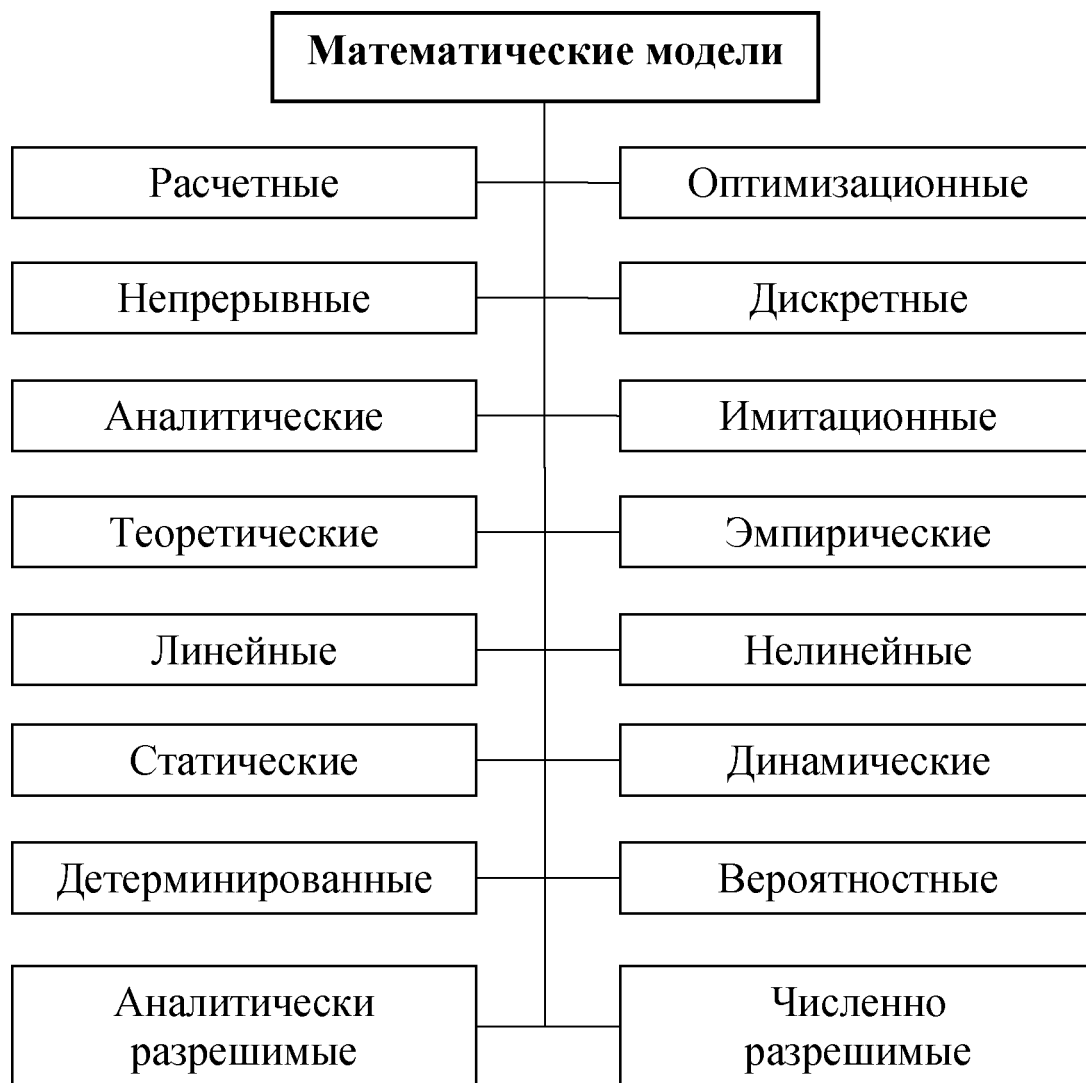


Рис. 13

1. *Расчетные* модели.
2. *Оптимизационные* модели.

Поясним их суть на простом примере. Пусть модель процесса описывается функцией вида:

$$y = f(x_1, x_2, \dots, x_n).$$

Эта зависимость будет *расчетной* моделью при решении следующей «прямой» задачи: «Вычислить значение выходной характеристики y процесса при заданных входных воздействиях x_1, x_2, \dots, x_n ».

При решении «обратной» задачи: «Определить значения воздействий x_1, x_2, \dots, x_n , при которых выходная характеристика y

рассматриваемого процесса достигает максимального (минимального) значения» приведенная выше зависимость является *оптимизационной* моделью вида:

$$f(x_1, x_2, \dots, x_n) \rightarrow \max (\min).$$

По природе моделируемых параметров и характеристик процессов на практике выделяются *непрерывные* и *дискретные* модели.

Математические модели по использованному в них формальному языку можно разбить на *аналитические* и *имитационные*.

Если связи между параметрами рассматриваемого процесса можно выразить в аналитической форме, т.е. в виде формул, то говорят об *аналитических* математических моделях.

Если какой-либо процесс не может быть описан в виде аналитической модели, его описывают с помощью специального алгоритма или программы. Такая модель называется *имитационной* моделью. По форме представления имитационная математическая модель является алгоритмической моделью, поскольку связь в ней между входными и выходными параметрами объекта (процесса) удастся описать лишь в форме алгоритма, пригодного для реализации в виде компьютерной программы. Имитационные модели – это обычно модели, в которых использован специальный язык моделирования или универсальный язык программирования.

По способу построения модели разбиваются на *теоретические* и *эмпирические* модели.

Теоретические модели отражают реальные структуры и процессы в исследуемых объектах, т.е. опираются на теорию их работы.

Эмпирические модели строятся на основе результатов экспериментов по изучению реакций объекта на изменение входных воздействий и условий окружающей среды. При этом теория работы объекта не рассматривается, сам объект представляет собой так называемый «черный ящик», а модель – некоторые интерполяционные или аппроксимирующие зависимости. Эмпирические модели могут быть построены на основе обработки экспериментальных данных, получаемых непосредственно на исследуемых объектах или с помощью их физических моделей.

По форме описания аналитические модели подразделяются на *линейные* и *нелинейные*. В линейной математической модели объекта (процесса) его параметры и характеристики связаны линейными соотношениями. Это означает, что при изменении какого-либо параметра процесса линейная модель предсказывает изменение зависящего от него выходного параметра, а при изменении двух или более параметров осуществляется сложение их влияний, т.е. такая математическая модель обладает свойством *суперпозиции*. Если модель не обладает свойством суперпозиции, то ее называют нелинейной.

Если все входящие в модель величины не зависят от времени, то имеем *статическую* модель объекта или процесса, в противном случае получаем *динамическую* модель. Статическое моделирование служит для описания состояния процесса в фиксированные моменты времени, а динамическое отражает поведение процесса во времени.

В *детерминированных* моделях все взаимосвязи, переменные и константы заданы точно, что приводит к однозначному определению получаемых с их помощью результатов.

Если часть или все параметры, входящие в модель по своей природе являются случайными величинами или случайными функциями, то модель относят к классу *вероятностных* моделей. В таких моделях задаются законы распределения используемых случайных величин и функций, что приводит к вероятностному характеру получаемых результатов.

Если аналитическое исследование процесса с использованием математических методов может быть доведено до уровня, когда результаты моделирования получены в виде формул, модели называются *аналитически разрешимыми*. В противном случае говорят о *численно разрешимых* аналитических моделях.

Аналитически разрешимые модели позволяют изучить свойства рассматриваемого процесса, применяя традиционные хорошо развитые методы классического математического анализа. Кроме того, знание аналитического выражения для выходных параметров позволяет с помощью формальных (математических) методов исследовать фундаментальные свойства моделируемого процесса и формировать рекомендации по его требуемому поведению. К сожалению, существующие в настоящее время математические методы позволяют получить аналитические решения лишь для относительно несложных математических моделей, хотя искусство моделирования состоит в построении именно таких моделей.

Численно разрешимые модели подразумевают использование методов вычислительной математики и компьютерной техники для получения требуемых результатов. При этом все исходные данные для моделирования и его результаты представляются в виде массивов чисел.

Формально математическую модель в общем виде можно представить выражением вида:

$$\mathfrak{M} = \{M, R_1, R_2, \dots, R_n\}. \quad (1)$$

Например, круг задается неравенством $x^2 + y^2 \leq a^2$, где x, y – множество точек на плоскости, a – действительное число. Тогда математическая модель (1) геометрического объекта «круг» выглядит следующим образом: $\mathfrak{M} = \{M, R_1\}$, где $M = \{x, y, a\}$, а $R_1 = \{x^2 + y^2 \leq a^2\}$.

Отметим дополнительно, что формальное описание объекта (процесса) должно быть таковым, чтобы модель удовлетворяла следующим свойствам:

1. Утверждения, полученные в рамках модели, могут быть проверены на истинность и ложность.
2. Утверждения, получаемые с помощью применяемой модели, должны обладать общностью.
3. Информативность утверждений в рамках формального описания должна быть больше, чем при отсутствии модели.

Входящие в выражение (1) отношения $R_i, i = \overline{1, n}$ должны описывать некоторые зависимости между элементами множества M . Эти зависимости в реальном моделировании отображают интересующие пользователя свойства объекта, процесса или явления. Такие отношения будем называть первичными отношениями.

Обязательным атрибутом применения любой модели должны являться некоторые правила вывода \mathcal{P} , позволяющие получать новые отношения между элементами множества M . Эти отношения (связи) Q_1, Q_2, \dots, Q_m должны нести в себе новые знания о моделируемом процессе, явлении или объекте. Таким образом, схема использования модели вида (1) записывается так:

$$\{R_1, R_2, \dots, R_n\} \xrightarrow{\mathcal{P}} \{Q_1, Q_2, \dots, Q_m\}. \quad (2)$$

Отметим, что правила вывода \mathcal{P} реализуются как всевозможные математические методы и приемы.

Например, при решении уравнения вида:

$$ax + b = 0$$

модель (1) конкретизируется как:

$$\mathfrak{M} = \{a, x, b, R_1\},$$

где

$$R_1: ax + b = 0.$$

Вторичное отношение при этом записывается в форме выражения:

$$Q_1: x = -\frac{b}{a}.$$

Правило его вывода \mathcal{P} получается следующим образом:

$$\{ax + b = 0\} \Rightarrow \{ax = -b\} \Rightarrow -\left\{x = -\frac{b}{a}\right\}.$$

Связь модели и реального мира представим схемой на Рис. 14.

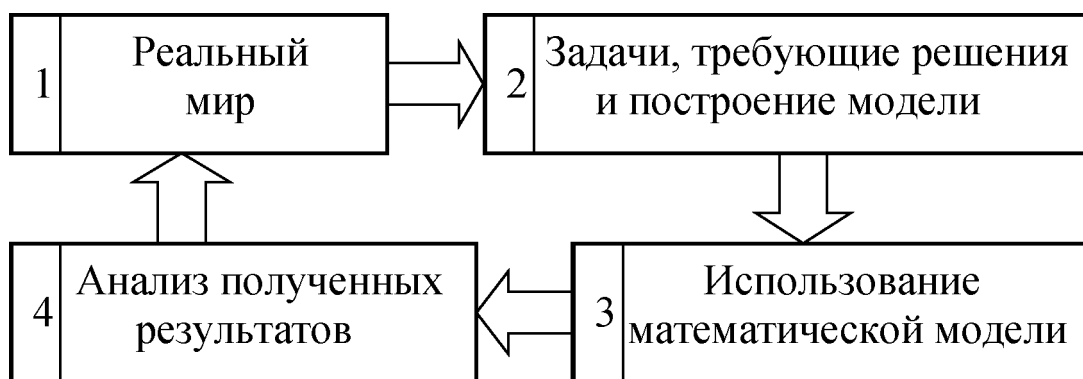


Рис. 14

Рассмотрим характеристику каждого этапа, представленного в этой схеме.

Этап 1. В реальной действительности возникает некоторая проблема и собирается вся относящаяся к ней информация. Эти данные используются для построения множества M и первичных отношений R_1, R_2, \dots, R_n .

Этап 2. На основе этой информации формируются соответствующие постановки задач, и строится модель вида (1).

Этап 3. С математической моделью проводятся математические операции (применяются правила вывода \mathcal{P}) и получают результаты в форме вторичных отношений (Q_1, Q_2, \dots, Q_m).

Этап 4. Эти результаты исследователь сопоставляет с реальным миром и делается вывод о том, получено ли решение сформированной задачи или нет.

Научный метод можно описать как циклический процесс, с помощью которого человечество получает новое знание из опыта. По мере накопления фактов, могут формироваться научные теории, которые всё более достоверно описывают объективную реальность.

Основной цикл научного метода можно разделить на 3 шага:

- Индукция;
- Дедукция;
- Верификация (проверка).

Индукция – это шаг, приводящий исследователя от наблюдений к образованию теорий. Теории бывают эмпирическими (в основном занимаются классификацией фактов, например, таблица Менделеева) и абстрактными (основываются на применении математической модели). Индуктивный шаг является творческим процессом, в основе которого лежит некоторая догадка, гипотеза, концепция.

Как только сформулирована теория, включающая модель вида (1), становится возможным применение аппарата математики (правила \mathcal{P}) для получения следствий из этой теории (Q_1, Q_2, \dots, Q_m). Это относится к дедукции, то есть к шагу, в котором получаются следствия теории, которые могут быть и неочевидными.

Как только с помощью теории получено некоторое количество следствий, их следует оценить с помощью экспериментальной проверки (верификация). При этом наряду с натурными, а часто и вместо них используются и вычислительные эксперименты.

При верификации возможны два случая:

- вновь полученные факты могут соответствовать уже известным результатам.
- для проверки результатов может потребоваться новая информация. В этом случае осуществляется возврат на этап 1 и цикл вновь повторяется.

Для любой теории третий этап (проверка) обязательно должен заканчиваться оценкой доверия (достоверности) построенной теории. Если достоверность низкая, то осуществляется разработка другой усовершенствованной теории, которая должна объяснять как старые факты, на которых базировалась отвергнутая теория, так и все новые факты, которые привели к ее отвержению.

Отметим основные особенности математического моделирования в прикладной информатике.

1. Информатика базируется на дискретных объектах и процессах. Поэтому в их моделях вида (1) в основном используется аппарат дискретной математики, хотя в некоторых задачах применяется аппарат «непрерывной» математики.

2. Объектами моделирования будем считать информационные системы и технологии, процессы их развития и эксплуатации, программные средства прикладной информатики, а также методы и средства защиты информации.

Далее на протяжении лекций Разд. II будет рассмотрен существующий и перспективный математический аппарат для моделирования дискретных объектов и процессов.

Раздел II. ЭЛЕМЕНТЫ СОВРЕМЕННОЙ ТЕОРИИ ОТНОШЕНИЙ ДИСКРЕТНЫХ МНОЖЕСТВ

Понятие отношения широко используется в теории множеств и функциональном анализе. Практически всегда это отношения непрерывных множеств с континуумом их элементов. Отношения дискретных множеств крайне редко встречаются в существующей литературе.

В данном разделе приводится систематическое изложение существующих подходов к описанию таких отношений, приводятся матричные операции над отношениями и предлагаются новые виды бинарных отношений для использования в моделях информационных систем и технологий.

5. Основы классической теории отношений дискретных множеств

Как было отмечено в Лекции №4, под математической моделью объекта, системы, процесса в общем случае понимается множество \mathfrak{M} , состоящее из совокупности элементов M произвольной природы, на которых определено конечное множество отношений (связей) R_1, R_2, \dots, R_n (см. выражение (1)).

Для построения связей между нечисловыми объектами используется формальное понятие дискретного отношения.

Понятие отношения множеств занимает значимое место во многих разделах современной математики, в которых в основном рассматриваются отношения непрерывных (континуальных) множеств, такие как отношения эквивалентности, предпочтения и порядка. Такой вид отношений широко используются в функциональном анализе, матлогике, теоретических основах векторной (многокритериальной) оптимизации и др.

Прикладным аспектам применения теории отношений посвящено незначительное число работ, в которых не указываются способы описания применяемых отношений, а их семантика задается субъективно исследователем или разработчиком соответствующих процессов и систем. Вместе с тем, в последнее время для моделирования сложных организационно-технических систем начинается использование дискретных множеств и отношений таких множеств. При этом для получения дополнительной информации о свойствах рассматриваемой системы используются известные виды операций над отношениями, входящими в ее теоретико-множественную модель.

Относительно скромные результаты применения такого подхода к моделированию и анализу сложных систем, объясняются, на наш взгляд, недостаточной информированностью исследователей и разработчиков таких систем о возможностях теории отношений дискретных множеств, особенно в части формального описания таких отношений и выполняемых над ними операций. Кроме этого, дополнительным сдерживающим фактором является отсутствие новых видов отношений дискретных множеств, позволяющих формализовать их семантику, используемую при решении различных прикладных задач.

Рассмотрим понятие и свойства классических отношений.

В основу понятия отношения полагается известное в математике понятие произведения множеств. Обозначим через $X = (x_1, x_2, \dots, x_n)$ упорядоченную последовательность n элементов, которую будем называть кортежем длины n . Примерами кортежей являются множество людей в очереди, множество слов во фразе, множество операторов в программе и т.п.

В отличие от множества, в кортеж могут входить одинаковые элементы (например, одинаковые операторы в программе).

Пусть задано n множеств X_1, X_2, \dots, X_n произвольной природы.

Определение 1. Множество всех кортежей (x_1, x_2, \dots, x_n) , таких, что $x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n$ называется прямым (декартовым) произведением множеств X_1, X_2, \dots, X_n и обозначается как $D = X_1 \times X_2 \times \dots \times X_n$.

Примером такого произведения является множество клеток шахматной доски:

$D = X_1 \times X_2 = \{(a, 1), (a, 2), \dots, (b, 1), (b, 2), \dots, (h, 8)\}$,
где $X_1 = \{a, b, c, d, e, f, g, h\}; X_2 = \{1, 2, 3, 4, 5, 6, 7, 8\}$,

Отметим, что перестановка «сомножителей» в произведении множеств даёт другое произведение, т.е. $X_1 \times X_2 \neq X_2 \times X_1$.

В приложениях часто декартово произведение строится для одинаковых множеств.

Определение 2. Отношением R на множествах X_1, X_2, \dots, X_n называется подмножество декартова произведения $X_1 \times X_2 \times \dots \times X_n$, которое обозначается как:

$$R \subseteq X_1 \times X_2 \times \dots \times X_n.$$

Другими словами, элементы кортежа (x_1, x_2, \dots, x_n) , где $x_i \in X_i, i = \overline{1, n}$ связаны отношением R тогда и только тогда, когда $(x_1, x_2, \dots, x_n) \in R$.

В зависимости от n различают бинарные отношения ($n = 2$), тернарные ($n = 3$) и другие виды отношений. Примером бинарного отношения является множество чёрных клеток шахматной доски.

В дальнейшем свойства и представления отношений будем рассматривать для бинарных отношений.

Пусть даны два произвольных множества X и Y . Декартовым произведением этих множеств, обозначаемым как:

$$D = X \times Y, \quad (3)$$

называется множество всех упорядоченных пар (x, y) элементов рассматриваемых множеств, где $x \in X$ и $y \in Y$.

Графическая иллюстрация произведения D для непрерывных (континуальных) множеств X и Y приведена на Рис. 15, из которого следует, что декартово произведение (3) формально описывает бесконечное множество точек с координатами (x, y) , лежащих в заштрихованном прямоугольнике.

Для случая дискретных множеств:

$$X = \{x_1, x_2, \dots, x_k\},$$

$$Y = \{y_1, y_2, \dots, y_m\}$$

графическое представление декартова произведения (3) приведено на Рис. 16.

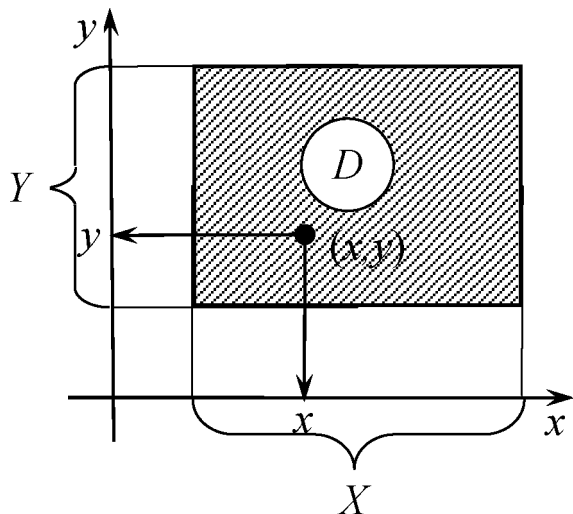


Рис. 15

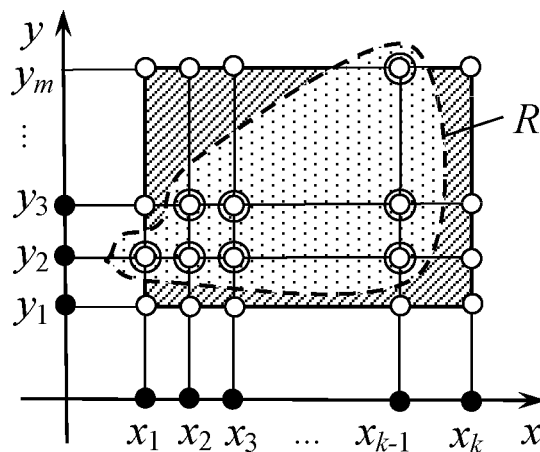


Рис. 16

В этом случае декартово произведение $D = X \times Y$ представляет собой полное дискретное множество упорядоченных пар (x_i, y_i) , где $x_i \in X$ и $y_i \in Y$.

Число таких пар, т.е. мощность множества D , представляющие собой плоскую «решетку» (см. Рис. 16), равно $|D| = km$.

С точки зрения теории графов, выражение (3) для дискретных множеств X и Y описывается полным двудольным ориентированным графом Γ^* , т.е. графом, в котором каждая вершина $x_i \in X$ связана ориентированными дугами со всеми вершинами, составляющими множество Y .

Пример.

Для множеств $X = \{x_1, x_2\}$ и $Y = \{y_1, y_2, y_3, y_4\}$ граф Γ^* приведен на Рис. 17.

Отметим, что составляющие отношения (3) необязательно должны иметь числовую природу. Например, множество X может описывать марки компьютеров, а Y - варианты их использования (настольный, носимый) и т.п.

С использованием понятия декартова произведения множеств вводится понятие бинарного отношения как некоторого подмножества множества D (см. Рис. 16), которое в некоторых случаях может совпадать с этим множеством.

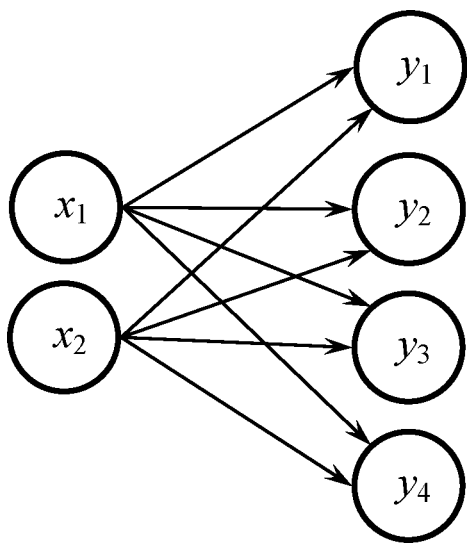


Рис. 17

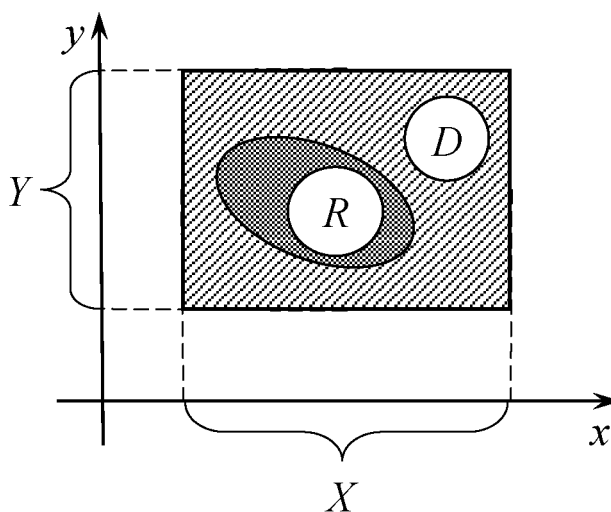


Рис. 18

Для декартова отношения (3) такое отношение записывается в следующей форме:

$$R \subseteq X \times Y, \quad (4)$$

где \subseteq означает включение множества R в множество D .

На Рис. 18. приведен геометрический смысл выражения (4) для множества D , представленного на Рис. 15.

Аналогичный смысл отношение (4) имеет и для дискретных множеств X и Y , описывая часть «решетки», представленной на Рис. 16.

В этом случае, отношение R описывает определенное множество упорядоченных пар (x_i, y_j) , где $i \in (\overline{1, k})$ и $j \in (\overline{1, m})$. Число таких пар носит название мощности отношения R и обозначается как $|R|$.

Рассмотрим основные свойства отношений для частного вида бинарного отношения:

$$R \subseteq A \times A,$$

которое основывается на одном множестве A .

На элементы отношений могут быть возложены дополнительные условия, которые отражают свойства рассматриваемых отношений.

1. Отношение R является рефлексивным отношением, если $(x R x)$ для всех $x \in A$.

Графически это иллюстрируется Рис. 19.

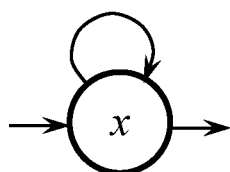


Рис. 19

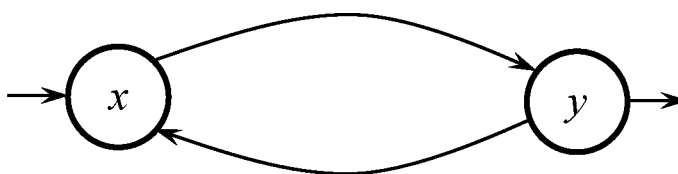


Рис. 20

2. Отношение R симметрично, если из $(x R y)$ следует, что $(y R x)$ (Рис. 20).

3. Отношение R транзитивно, если из $(x R y)$ и $(y R z)$ следует, что $(x R z)$ (Рис. 21).

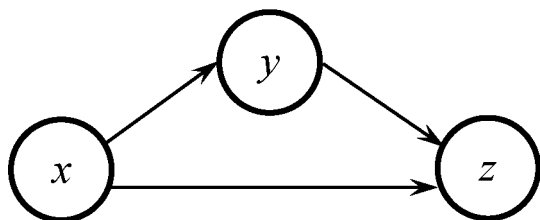


Рис. 21

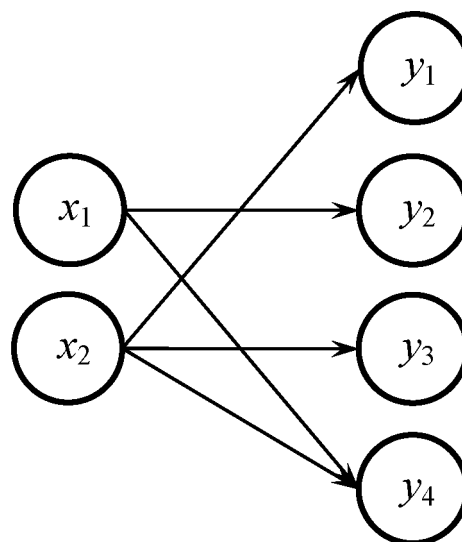


Рис. 22

4. Отношение R антисимметрично, если из $(x R y)$ и $(y R x)$ следует, что $x = y$.

Анализ отмеченных выше свойств отношений показывает, что они могут быть использованы при решении задач анализа структуры и информационных связей в сложных системах.

В существующих приложениях выделяют три вида отношений:

- Отношение эквивалентности, когда любой элемент множества при конкретном рассмотрении может быть заменен другим элементом. Например, отношение «быть на первом курсе» на множестве студентов факультета, отношение «подобия» на множестве треугольников.

Для отношения эквивалентности выполняются свойства симметричности, транзитивности, рефлексивности.

- Отношение порядка («раньше», «позже», «меньше», «больше» и т.п.). Для числовых множеств выделяются отношения нестрогого порядка (\leq) и отношения строгого порядка ($<$).

Отношение нестрогого порядка рефлексивно, антисимметрично, транзитивно.

Отношение строгого порядка антирефлексивно, несимметрично, транзитивно.

- Отношение доминирования, которое используется в основном для множеств, имеющих нечисловую природу.

Будем говорить, что элемент x доминирует над y , если x в чем-то превосходит y . Например, команда « x » сильнее команды « y », или лицо « x » пользуется авторитетом у лица « y ».

Доминирование элементов обозначается как $x \succ y$.

Определение 3. В множестве A введено отношение доминирования, если его элементы обладают следующими свойствами:

- Никакой элемент множества не может доминировать себя (антирефлексивность).
- В любой паре элементов в точности один элемент доминирует второго (несимметричность).
- Транзитивность отсутствует.

Формально отношение (4) для дискретных множеств X и Y описывается матрицей $A = [a_{ij}]$ размерности $k \times m$ с элементами вида:

$$a_{ij} = \begin{cases} 1, & \text{если пара } (x_i, y_j) \in R; \\ 0, & \text{если пара } (x_i, y_j) \notin R, \end{cases} \quad (5)$$

где $i = 1, 2, \dots, k$, $j = 1, 2, \dots, m$.

Если в матрице A присутствует i -я нулевая строка или j -й нулевой столбец, то это означает, что определенные элементы $x_i \in X$ или $y_j \in Y$ не участвуют в формировании отношения (4).

Матрица A с точки зрения теории графов может рассматриваться как матрица смежности двудольного ориентирован-

ного графа, вершины которого описываются элементами множеств X и Y .

Такая матрица описывает факты наличия или отсутствия в графе ориентированных дуг, связывающих определенные вершины $x_i \in X$ с вершинами $y_j \in Y$.

Пример.

Пусть для множеств $X = \{x_1, x_2\}$ и $Y = \{y_1, y_2, y_3, y_4\}$ отношение (4) описывается матрицей вида:

$$A = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{matrix}.$$

Из смысла ее элементов, определяемого выражением (5), следует, что отношение R включает в себя такие пары элементов множеств X и Y , как (x_1, y_2) ; (x_1, y_4) ; (x_2, y_1) ; (x_2, y_3) ; (x_2, y_4) .

Соответствующий этому отношению ориентированный граф, представлен на Рис. 22.

Из сравнения с графом, приведенным на Рис. 17, следует, что в полученном графе отсутствует часть дуг. Отметим, что матрица A для декартова произведения D , представленного на этом рисунке, имеет вид:

$$A = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{matrix}.$$

Если отношение R строится на основе трех множеств X, Y, Z , то такое отношение называется тернарным и обозначается как

$$Q \subseteq X \times Y \times Z. \quad (6)$$

Отношение (6) для дискретных множеств так же описывается матрицей A с элементами:

$$a_{ijr} = \begin{cases} 1, & \text{если пара } (x_i, y_i, z_r) \in R; \\ 0, & \text{в противном случае,} \end{cases} \quad (7)$$

где $x_i \in X, y_j \in Y, z_r \in Z$.

Представление отношения (6) в виде трехдольного ориентированного графа (орграфа) приведено на Рис. 23.

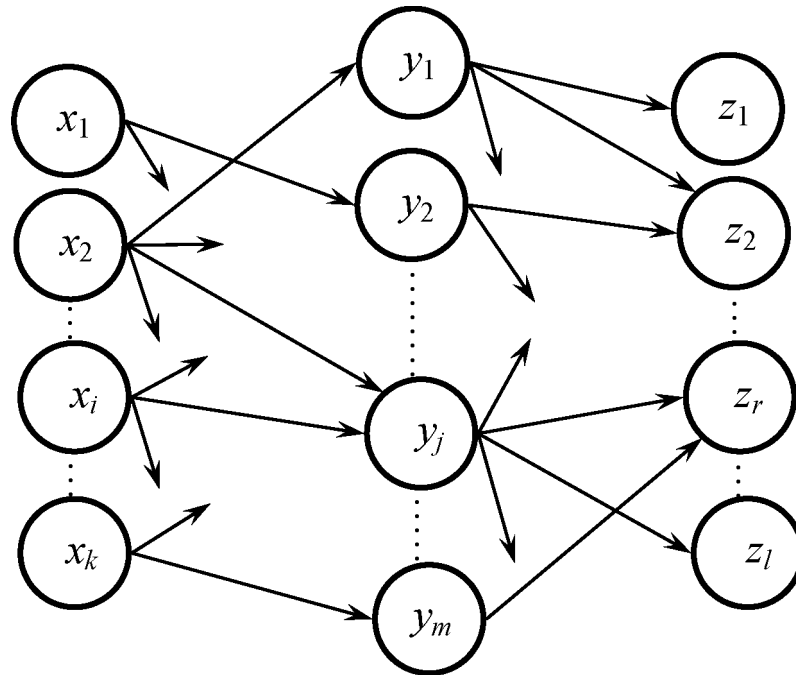


Рис. 23

В общем случае n -арные отношения вида:

$$Q \subseteq X_1 \times X_2 \times \dots \times X_n. \quad (8)$$

формально может быть описано матрицей A с n -индексными элементами, принимающими значения 0 или 1. Графическое представление отношения (8) изображается n -дольным орграфом.

Другие свойства и примеры бинарных отношений приведены в книге [Кук Д., Бейз Г.].

Отметим, что рассмотренные выше виды отношений, согласно классификации, приведенной на Рис. 13, используются для построения расчетных дискретных статических детерминированных математических моделей в прикладной информатике.

6. Теоретико-множественные операции над отношениями

Рассмотрим существующие операции над отношениями и их формальные представления для компьютерных реализаций. Отметим, что последнее отсутствует в доступной литературе.

Из Лекции №5 следует, что любые отношения по своей природе представляются некоторыми множествами. Это означает, что над ними можно производить как известные теоретико-множественные операции (объединение и пересечение множеств и др.), так и специальные операции (инверсию отношения, композицию отношений и др.).

Отметим, что в существующей литературе приведен только смысл некоторых из таких операций без указания методов получения результирующих отношений. Рассмотрим смысл операций объединения и пересечения отношений.

Пусть даны два бинарных отношения:

$$Q_1 \subseteq X \times Y; Q_2 \subseteq X \times Y, \quad (9)$$

определенных на декартовом произведении (4). Графическое представление этих отношений для непрерывных множеств X и Y (см. Рис.18) приведено на Рис. 24.

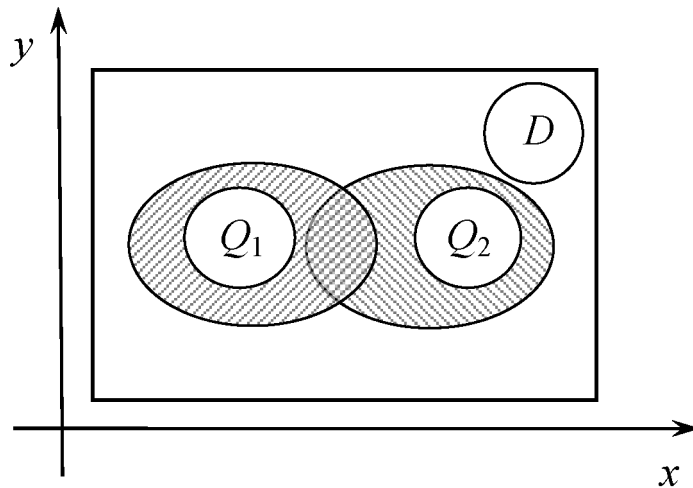


Рис. 24

Тогда объединение отношений (9), которое записывается как:

$$R_1 = Q_1 \cup Q_2, \quad (10)$$

представляет собой множество упорядоченных пар (x, y) , где $x \in X, y \in Y$, которые принадлежат отношениям Q_1 и Q_2 сразу или по отдельности.

Пересечение отношений Q_1 и Q_2 представляется отношением вида:

$$R_2 = Q_1 \cap Q_2. \quad (11)$$

Это отношение определяется как множество упорядоченных пар (x, y) , где $x \in X, y \in Y$, общих для отношений Q_1 и Q_2 .

На Рис. 24 областью существования отношения (11) является множество пар (x, y) , лежащих в геометрическом месте точек, обозначенном двойной штриховкой.

Рассмотрим способы формального описания отношений (10) и (11) для дискретных множеств $X = \{x_1, x_2, \dots, x_k\}$ и $Y = \{y_1, y_2, \dots, y_m\}$.

Как было отмечено, любое бинарное отношение (4), построенное на таких множествах, описывается матрицей A , элементы которой имеют вид, определяемый выражением (5).

Будем считать, что отношения (9) описываются матрицами A_1 и A_2 одинаковой размерности $k \times m$.

Тогда матрица A , описывающая отношение (10), может быть получена как:

$$A = A_1 \oplus A_2, \quad (12)$$

где \oplus – операция логического поэлементного сложения элементов матриц A_1 и A_2 , которое выполняется по следующим правилам:

$$0 \oplus 0 = 0; \quad 0 \oplus 1 = 1; \quad 1 \oplus 0 = 1; \quad 1 \oplus 1 = 1. \quad (13)$$

Матрица A , описывающая отношения (11), получается по формуле вида:

$$A = A_1 \otimes A_2, \quad (14)$$

где \otimes - операция логического поэлементного умножения, выполняемая с использованием следующих правил:

$$0 \otimes 0 = 0; \quad 0 \otimes 1 = 0; \quad 1 \otimes 0 = 0; \quad 1 \otimes 1 = 1. \quad (15)$$

Пример.

Проиллюстрируем выполнение операций объединения и пересечения отношений, используя в качестве отношения Q_1 отношение из предыдущего примера, матричное представление которого имеет вид:

$$A_1 = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{matrix}.$$

Отношение Q_2 будем описывать орграфом, приведенным на Рис. 25.

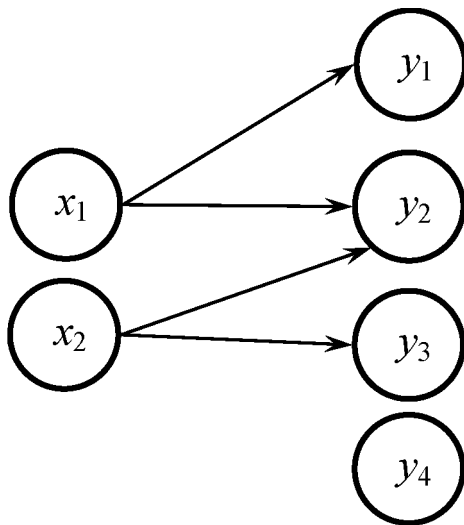


Рис. 25

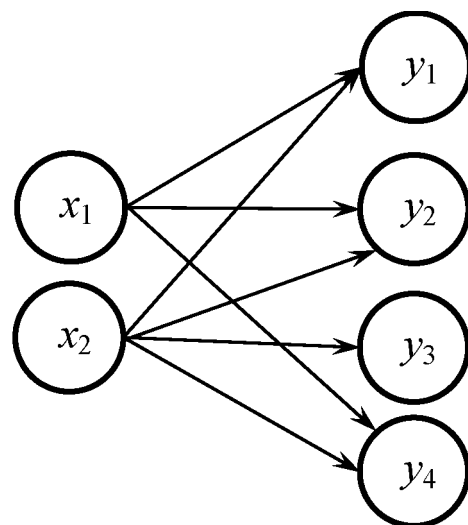


Рис. 26

Матрица A_2 , описывающая отношение Q_2 , записывается как:

$$A_2 = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}.$$

Используя выражения (12) и (13), построим матрицу A , определяющую объединение рассматриваемых отношений:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \oplus \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Орграф, построенный на основе полученной матрицы A , представлен на Рис. 26.

Из анализа графов, приведенных на Рис. 22 и Рис. 25, следует, что в результате объединения отношений добавились пары (x_1, y_1) ; (x_2, y_2) , которые отсутствовали в отношении Q_1 .

Пересечение отношений Q_1 и Q_2 формально описывается матрицей A , которая строится на основе выражений (13) и (15) как:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Графическое представление этого отношения приведено на Рис. 27, из которого следует, что общими для отношений Q_1 и Q_2 являются пары (x_1, y_2) ; (x_2, y_3) .

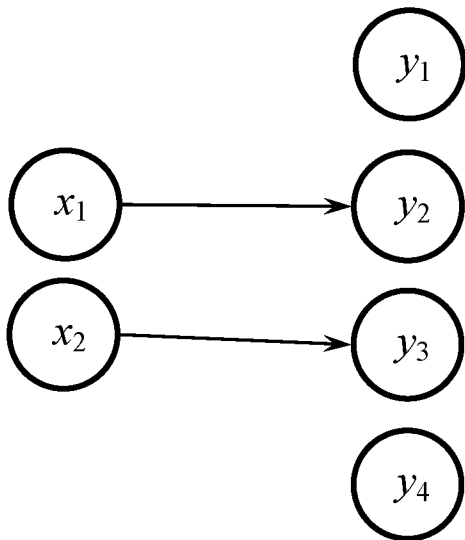


Рис. 27

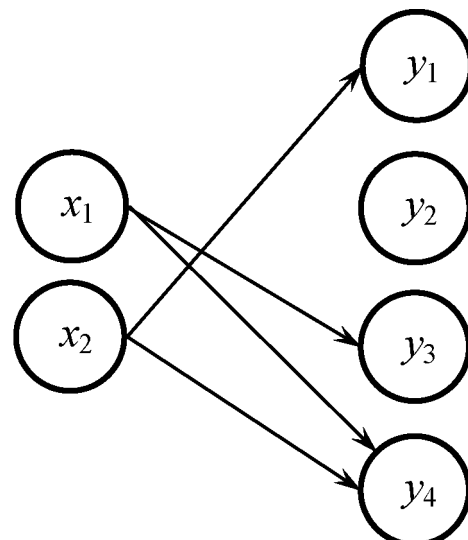


Рис. 28

Отметим, что если в результате применения операции (15) к матрицам A_1 и A_2 будет получена матрица A с элементами $a_{ij} = 0$ для всех $i = (\overline{1, k})$ и $j = (\overline{1, m})$, то можно утверждать,

что отношения Q_1 и Q_2 не имеют общих пар (x_i, y_j) , т.е. R_2 будет пустым множеством ($R_2 = \emptyset$).

Если заданы N бинарных отношений вида:

$$Q_1 \subseteq X \times Y, Q_2 \subseteq X \times Y, \dots, Q_n \subseteq X \times Y, \quad (16)$$

то операция объединения этих отношений записывается как:

$$R_1 = \bigcup_{s=1}^N Q_s. \quad (17)$$

Операцию пересечения отношений (16) можно представить в следующем виде:

$$R_2 = \bigcap_{s=1}^N Q_s. \quad (18)$$

Матричные представления отношений (17) и (18) можно получить при обобщении выражений (12) и (14). В частности, результирующая матрица A , описывающая отношение (17), получается как:

$$A = A_1 \oplus A_2 \oplus \dots \oplus A_N$$

Для отношения (18) такая матрица вычисляется с помощью выражения вида:

$$A = A_1 \otimes A_2 \otimes \dots \otimes A_N$$

В этих формулах матрицы A_s описывают отношения Q_s , $s = \overline{1, N}$, входящие в выражение (16).

Операцию разности отношений Q_1 и Q_2 , описываемых выражениями (9), будем обозначать следующим образом:

$$R_3 = Q_1 \setminus Q_2. \quad (19)$$

Следуя смыслу соответствующей теоретико-множественной операции в состав отношения R_3 должны войти пары (x_i, y_j) , $x_i \in X, y_j \in Y$, которые составляют отношение Q_1 за

исключением пар, принадлежащих отношению R_2 , которое описывается выражением (11).

Согласно Рис. 24 отношение (19) образуют точки множества Q_1 , из которого исключены точки, входящие в область, обозначенную двойной штриховкой.

Матрицу A , описывающую отношение R_3 , будем вычислять по формуле вида:

$$A = A_1 \ominus A_2, \quad (20)$$

где A_1, A_2 - матричные описания отношений Q_1 и Q_2 ; \ominus - операция логической поэлементной разности матриц, выполняемая по следующим правилам:

$$0 \ominus 0 = 0; \quad 0 \ominus 1 = 0; \quad 1 \ominus 0 = 1; \quad 1 \ominus 1 = 0. \quad (21)$$

Рассмотрим пример построения отношения (19).

Пример.

Выберем в качестве отношения Q_1 бинарное отношение из примера, представленного на Рис. 22. Отношение Q_2 в форме орграфа приведено на Рис. 25.

Используя выражения (20) и (21), построим матрицу A , определяющую разность отношений Q_1 и Q_2 :

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \ominus \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Орграф, построенный на основе полученной матрицы A , представлен на Рис. 28.

Из анализа графа, приведенного на этом рисунке, следует, что в результате применения операции «вычитания» отношений остались пары (x_1, y_3) ; (x_1, y_4) ; (x_2, y_1) ; (x_2, y_4) .

Операцию построения дополнения заданного отношения $Q_1 \subseteq X \times Y$ определим как операцию формирования отношения вида:

$$R_4 = D \setminus Q_1, \quad (22)$$

где D - декартовое произведение множеств X и Y , описываемое выражением вида (3).

Согласно Рис. 24, отношение R_4 составляют пары $(x, y) \notin Q$ при $x \in X$; $y \in Y$.

Матричное представление отношения (22) получается с использованием формулы вида:

$$A = A_D - A_1, \quad (23)$$

где A_D - матрица, описывающая множество D , представленное на Рис. 16, элементами которой являются единицы; A_1 - матрица, описывающая заданное отношение Q_1 .

В формуле (23) в связи с составом входящих в нее матриц использована общепринятая в матричной алгебре операция вычитания матриц.

Пример.

Проиллюстрируем выполнение операции дополнения отношения, используя отношение Q_1 , матричное представление которого имеет вид:

$$A_1 = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{matrix}.$$

Применяя выражение (23), построим матрицу A , определяющую дополнение этого отношения:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Орграф, построенный на основе полученной матрицы A представлен на Рис. 29.

Из анализа графов, приведенных на Рис. 22 и Рис. 29 следует, что в результате построения дополнения отношения Q_1 были сформированы пары (x_1, y_1) ; (x_2, y_2) ; (x_1, y_3) , которые отсутствовали в отношении Q_1 .

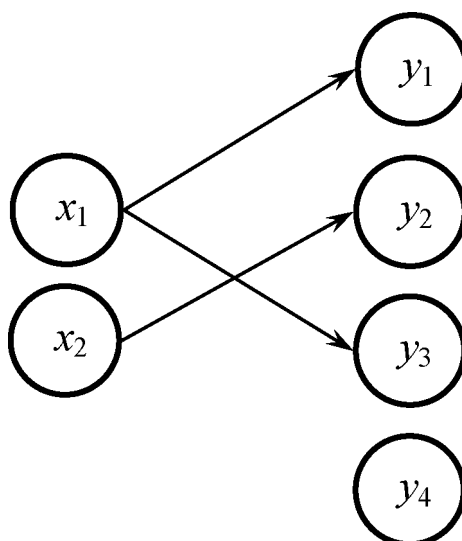


Рис. 29

При большой размерности матриц, описывающих отношения, над которыми выполняются соответствующие операции, можно использовать «списочное» представление соответствующих им орграфов. Например, граф, приведенный на Рис. 29, описывается списком вида:

$$x_1(y_1, y_3); \quad x_2(y_2).$$

Самостоятельная работа.

Разработать для «списочного» представления отношений алгоритмы выполнения операций их объединения, пересечения, вычитания и дополнения.

7. Специальные операции над отношениями

Рассмотрим смысл и представления этой группы операций.

Инверсией отношения R , представленного выражением (4), называется построение обратного отношения вида:

$$R^{-1} \subseteq Y \times X. \quad (24)$$

Это отношение образуется подмножеством упорядоченных пар (y_j, x_i) , $x_i \in X, y_j \in Y$, входящих в декартово произведение (3). Формально отношение R^{-1} представляется транс-

понированной матрицей A^T , полученной на основе матрицы A , описывающей исходное отношение R .

Пример.

Для отношения, представленного на Рис. 22, матричное и графическое представления отношения R^{-1} приведено на Рис. 30.

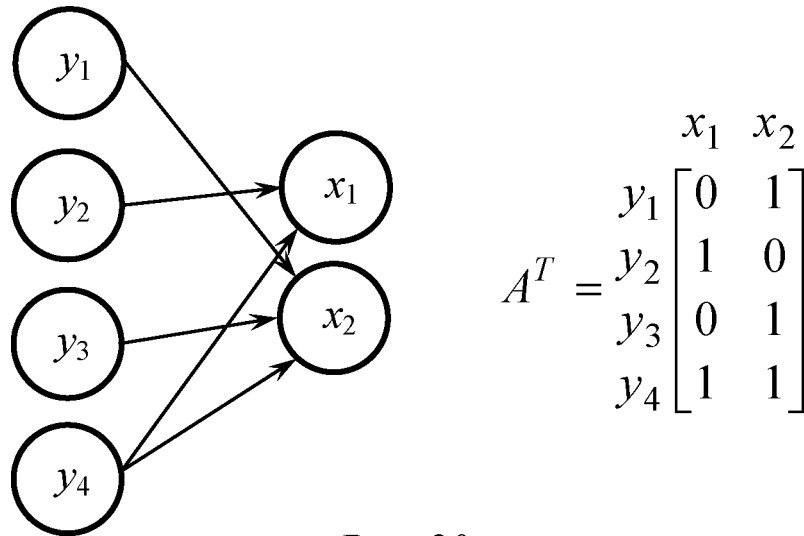


Рис. 30

В общем случае отношение, обратное к n -арному отношению (8), записывается как:

$$R^{-1} \subseteq X_n \times X_{n-1} \times \dots \times X_1.$$

Пусть даны два отношения вида:

$$Q_1 \subseteq X \times Y; Q_2 \subseteq Y \times Z. \quad (25)$$

Тогда композицией этих отношений называется построение следующего отношения:

$$R = Q_1 \circ Q_2 = (X \times Y) \circ (Y \times Z) \subseteq X \times Z. \quad (26)$$

Смысл операции композиции состоит в построении отношения над множествами X и Z с исключением общего для отношений (25) множества Y , но с учетом связей между множествами X , Y и Y , Z .

Пример.

Выберем в качестве отношения Q_1 бинарное отношение, представленное на Рис. 22. Отношение Q_2 в форме орграфа и

матрицы A_2 представлено на Рис. 31. Из указанных рисунков следует, что отношения Q_1 и Q_2 определяются как:

$$\begin{aligned} Q_1 &= \{(x_1, y_2)(x_1, y_4); (x_2, y_1); (x_2, y_3); (x_2, y_4)\}, \\ Q_2 &= \{(y_1, z_1); (y_2, z_1); (y_3, z_2); (y_3, z_3); (y_4, z_3)\}. \end{aligned} \quad (27)$$

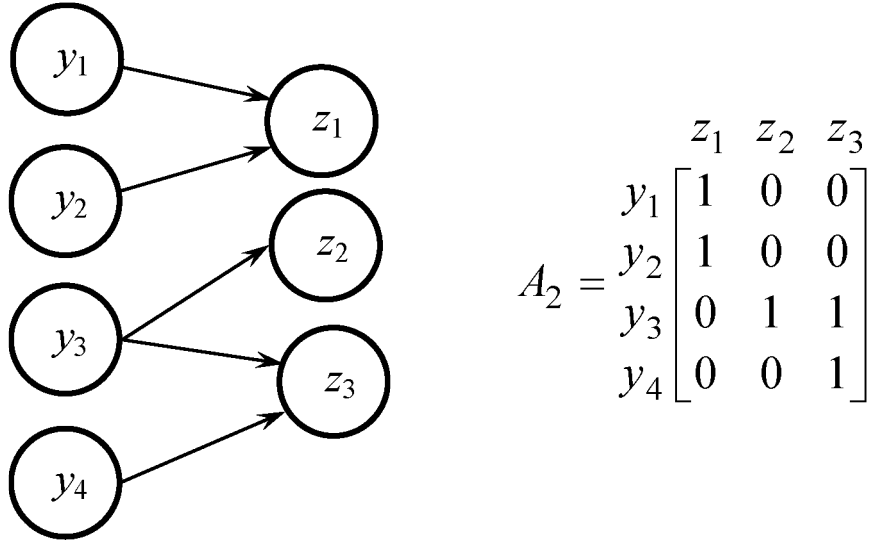


Рис. 31

Композиция отношений (26) строится на основе анализа пар $(x_i, y_j) \in Q_1$ и $(y_j, z_r) \in Q_2$ по следующему решающему правилу:

Если в некоторой из пар, входящих в множества (27) встречается одинаковый элемент $y_j \in Y$, то формируется пара $(x_i, z_r) \in R$. При этом, повторяющиеся пары вида (x_i, z_r) включаются в отношение R только один раз.

Применим это решающее правило к отношениям (27).

Элемент $y_1 \in Y$ входит в пары $(x_2, y_1) \in Q_1$ и $(y_1, z_1) \in Q_2$. Это означает, что в отношении R должны быть включена пара (x_2, z_1) , т.е. $(x_2, z_1) \in R$. Общим для пар $(x_2, y_1) \in Q_1$ и $(y_1, z_1) \in Q_2$ является элемент $y_2 \in Y$. Отсюда следует, что пара $(x_1, z_1) \in R$. Элемент $y_3 \in Y$ является общим для пар $(x_2, y_3) \in Q_1$ и $(y_3, z_2) \in Q_2$; $(y_3, z_3) \in Q_2$. Это означает, что в формируемое отношение включаются пары $(x_2, z_2) \in R$ и

$(x_2, z_3) \in R$. Общим для пар $(x_1, y_4) \in Q_1$, $(x_2, y_4) \in Q_1$ и $(y_4, z_3) \in Q_2$ является элемент $y_4 \in Y$, откуда следует, что $(x_1, z_3) \in R$ и $(x_2, z_3) \in R$.

Таким образом, композиция отношений, описываемых выражениями (27), определяется как множество вида:

$$R = \{(x_2, z_1); (x_1, z_1); (x_2, z_2); (x_2, z_3); (x_1, z_3)\}.$$

Графическое представление полученного отношения и соответствующая ему матрица A приведены на Рис. 32.

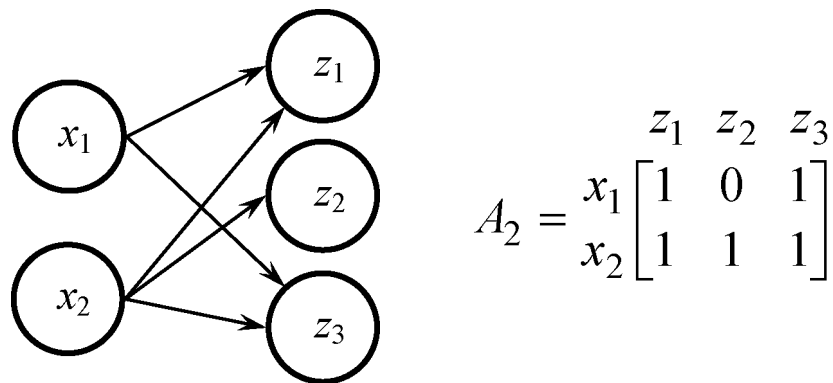


Рис. 32

Отметим, что операция композиции может быть применена к любому числу отношений. Например, композиция отношений:

$$Q_1 \subseteq X \times Y, \quad Q_2 \subseteq Y \times Z, \quad Q_3 \subseteq Z \times W$$

будет иметь вид:

$$R = Q_1 \circ Q_2 \circ Q_3 \subseteq X \times W.$$

Приведенное решающее правило для построения композиции бинарных отношений можно более просто реализовать с помощью специальных операций с матрицами A_1 и A_2 , описываемыми отношениями (15).

Пусть отношения (25) строятся на множествах $X = \{x_1, x_2, \dots, x_k\}$; $Y = \{y_1, y_2, \dots, y_m\}$ и $Z = \{z_1, z_2, \dots, z_n\}$.

Отсюда следует, что матрицы A_1 и A_2 имеют соответственно размерности $(k \times m)$ и $(m \times n)$. Для построения мат-

рицы A , описывающей результирующее отношение (26), будем использовать существующее правило умножения прямоугольных матриц с учетом логического умножения (15) их элементов.

Сформируем на основе матриц A_1 и A_2 вспомогательную матрицу $C = [c_{ir}]$ размерности $(k \times n)$. Элементы этой матрицы предлагается вычислять по формуле вида:

$$c_{ir} = \sum_{j=1}^m \left(a_{ij}^{(1)} \otimes a_{jr}^{(2)} \right), \quad i = (\overline{1, k}), r = (\overline{1, n}). \quad (28)$$

Эти величины определяют число элементов множества Y , с помощью которых осуществляется связь элементов множества X с элементами множества Y .

Тогда элементы матрицы $A = [a_{ir}]$, описывающей отношение $R \subseteq Q_1 \circ Q_2$, определяются как:

$$a_{ir} = \begin{cases} 1, & \text{если } c_{ir} > 0; \\ 0, & \text{если } c_{ir} = 0. \end{cases} \quad (29)$$

где $i = 1, 2, \dots, k$ и $r = 1, 2, \dots, n$.

Пример.

Проиллюстрируем полученные формулы на исходных данных, которые имеют вид:

$$A_1 = \begin{matrix} & \begin{matrix} y_1 & y_2 & y_3 & y_4 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{matrix}; \quad A_2 = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 \end{matrix} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

В этом примере имеем, что $k = 2$; $m = 4$ и $n = 3$. Отсюда следует, что вспомогательная матрица C будет иметь размерность 2×3 . Элементы этой матрицы будем вычислять по формуле (28) при $m = 4$, $i = (\overline{1, 2})$, $r = (\overline{1, 3})$.

Проводя логическое умножение элементов соответствующей строки матрицы A_1 на элементы соответствующего столбца матрицы A_2 , получим следующие значения величин c_{ir} :

$$c_{11} = (0 \otimes 1) + (1 \otimes 1) + (0 \otimes 0) + (1 \otimes 0) = 1;$$

$$c_{12} = (0 \otimes 0) + (1 \otimes 0) + (0 \otimes 1) + (1 \otimes 0) = 0;$$

$$c_{13} = (0 \otimes 0) + (1 \otimes 0) + (0 \otimes 1) + (1 \otimes 1) = 1;$$

$$c_{21} = (1 \otimes 1) + (0 \otimes 1) + (1 \otimes 0) + (1 \otimes 0) = 1;$$

$$c_{22} = (1 \otimes 0) + (0 \otimes 0) + (1 \otimes 1) + (1 \otimes 0) = 1;$$

$$c_{23} = (0 \otimes 1) + (0 \otimes 0) + (1 \otimes 1) + (1 \otimes 1) = 2.$$

Из анализа этих значений следует, что элементы $x_1 \in X$ и $z_2 \in Z$ не имеют связи между собой из-за отсутствия общих элементов из множества Y . Элементы x_1 и x_2 связаны с элементами z_1, z_2, z_3 с помощью одного соответствующего элемента множества Y . Элементы $x_2 \in X$ и $z_3 \in Z$ имеют два общих элемента в этом множестве.

Используя приведенное правило построения значений результирующей матрицы A , получаем ее вид и графическое представление искомого отношения R , приведенное на Рис. 32.

Введем в рассмотрение специальную операцию соединения отношений вида (25). Смысл этой операции состоит в формировании на их основе следующего отношения:

$$R = Q_1 \bar{\cup} Q_2 = (X \times Y) \bar{\cup} (Y \times Z) \subseteq X \times Y \times Z, \quad (30)$$

Из этого выражения следует, что такая операция формируется как тернарное отношение вида (6) на основе общего для двух бинарных отношений Q_1 и Q_2 множества Y .

Матрица $A = [a_{ijr}]$, описывающая отношение (30), строится по следующему решающему правилу:

Значение элемента a_{ijr} этой матрицы полагается равным единице, если в матрицах $A_1 = [a_{ij}^{(1)}]$ и $A_2 = [a_{jr}^{(2)}]$, опи-

сывающих отношения Q_1 и Q_2 , существуют элементы $a_{ij}^{(1)} = 1$ и $a_{jr}^{(2)} = 1$. Если хотя бы одно из этих значений равно нулю, то соответствующий элемент $a_{ijr} = 0$.

С использованием правил логического умножения вида (15) элементы матрицы A могут быть вычислены по формуле вида:

$$a_{ijr} = a_{ij}^{(1)} \otimes a_{jr}^{(2)}, i = (\overline{1, k}), j = (\overline{1, m}), r = (\overline{1, l}). \quad (31)$$

Отметим, что операция соединения может быть применена к любому числу отношений, имеющему попарно одинаковые множества.

Например, соединение рассмотренных выше отношений Q_1 , Q_2 и Q_3 будет иметь вид:

$$R = Q_1 \bar{\cup} Q_2 \bar{\cup} Q_3 \subseteq X \times Y \times Z \times W. \quad (32)$$

В общем случае операция соединения произвольного числа отношений описывается формулой вида:

$$R = \bigcup_{s=1}^N Q_s. \quad (33)$$

Результирующая N -индексная матрица отношения R может быть сформирована по формуле, аналогичной формуле (31).

Пример.

Пусть отношения Q_1 и Q_2 , представленные орграфами, изображенными на Рис. 19 и Рис. 28, описываются матрицами:

$$A_1 = [a_{ij}^{(1)}] = \begin{matrix} & \begin{matrix} y_1 & y_2 & y_3 & y_4 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \end{matrix}; \quad A_2 = [a_{jr}^{(2)}] = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 \end{matrix} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

Тогда, используя формулу (31) и перебирая всевозможные сочетания индексов i, j, r из интервалов $i = (\overline{1, 2})$, $j = (\overline{1, 4})$, $r = (\overline{1, 3})$, получаем следующие ненулевые элементы матрицы $A = [a_{ijr}]$ размерности $2 \times 4 \times 3$:

$$a_{121} = 1; \quad a_{143} = 1; \quad a_{211} = 1; \quad a_{232} = 1; \quad a_{233} = 1; \quad a_{243} = 1.$$

Орграф отношения R , построенный на основе этих значений, приведен на Рис. 33.

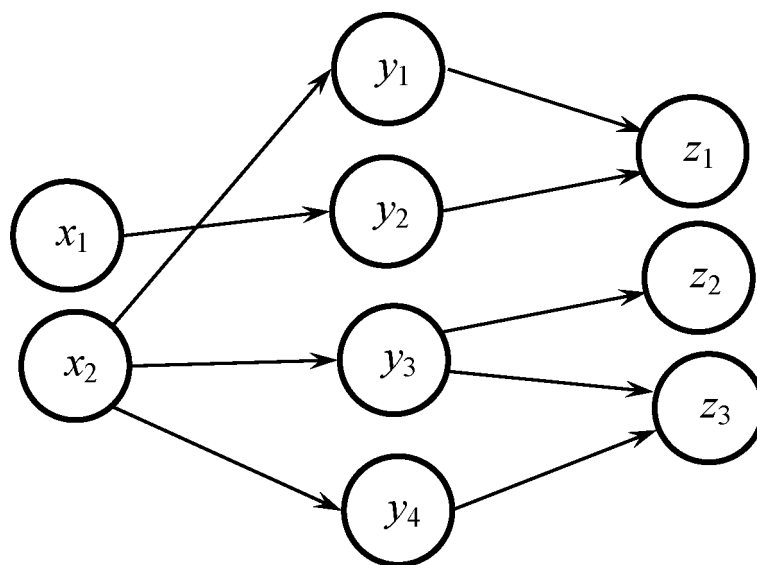


Рис. 33

Из этого рисунка следует достоверность выражения (31) и предлагаемой методики формирования соединения бинарных отношений.

Самостоятельная работа.

Разработать алгоритмы реализации специальных операций над отношениями, заданными «списочным» методом.

8. Случайные бинарные отношения

Рассмотренные выше отношения описывали фиксированный (детерминированный) состав множеств X и Y . Для представления вероятностных моделей (см. Рис. 13) необходимо описать ситуации, когда такие множества имеют случайный

характер. Такой ситуации соответствуют ориентированные графы со случайным составом их вершин.

Определение 4.

Случайным бинарным отношением на множествах $X = \{x_1, x_2, \dots, x_n\}$ и $Y = \{y_1, y_2, \dots, y_m\}$, законы распределения которых имеют вид:

$$\begin{aligned} P\{x_i \in X\} &= p_i; \\ P\{y_j \in Y\} &= q_j, \\ i &= (\overline{1, n}), \quad j = (\overline{1, m}), \end{aligned}$$

будем называть отношение:

$$\mathcal{R} \subseteq X \times Y, \quad (34)$$

Укажем некоторые свойства этого отношения.

Потенциальные связи между элементами множеств X и Y формально описываются матрицей $A = [a_{ij}]$. Введенные выше вероятности p_i и q_j должны удовлетворять условиям:

$$\begin{aligned} \sum_{i=1}^n p_i &= 1; \\ \sum_{j=1}^m q_j &= 1. \end{aligned}$$

Вероятность связи элементов $x_i \in X$ и $y_j \in Y$ представляется формулой вида:

$$P_{ij} = a_{ij} p_i q_j, \quad i = (\overline{1, n}), \quad j = (\overline{1, m}).$$

Таким образом, формальным представлением отношения (34) будем считать матрицу:

$$\mathcal{P} = [P_{ij}]_{n \times m}.$$

Рассмотрим один из подходов к формированию реализаций отношения \mathcal{R} .

Введем в рассмотрение для каждой i и j случайные события B_i и C_j , которые имеют следующие вероятности их наступления:

$$\begin{aligned}
P(B_i) &= p_i; \\
P(C_j) &= q_j, \\
i &= (\overline{1, n}), \quad j = (\overline{1, m}).
\end{aligned}$$

Тогда для формирования s -й реализации \mathcal{R}_s отношения (34) с использованием метода Монте-Карло [Вентцель Е.С.] в N экспериментах будем применять равномерно распределенные случайные числа $\alpha_k^{(s)} \in (0, 1)$ и $\beta_k^{(s)} \in (0, 1)$, $k = (\overline{1, N})$, $s = (\overline{1, S})$, которые формируются с помощью датчика таких случайных чисел (ДСЧ). Примем, что число генераций применяемых случайных чисел задается как: $N = \max\{n, m\}$.

Генерацию и отображение N реализаций отношения (34) будем проводить по следующему алгоритму:

1. Полагаем $s = 1$.
2. Полагаем $i = 1$.
3. Полагаем $j = 1$.
4. Полагаем $k = 1$.
5. Формируем с помощью датчика случайных чисел значения $\alpha_k^{(s)}$ и $\beta_k^{(s)}$.
6. Проверяем одновременное выполнение условий:

$$\begin{aligned}
\sum_{k=1}^{i-1} p_k &< \alpha_k^{(s)} < \sum_{k=1}^i p_k; \\
\sum_{r=1}^{j-1} q_r &< \beta_k^{(s)} < \sum_{r=1}^j q_r.
\end{aligned}$$

При их выполнении пара (x_i, y_j) включается в состав реализации \mathcal{R}_s , что представляется значениями $a_{ij}^{(s)} = 1$. При нарушении хотя бы одного неравенства, полагаем $a_{ij}^{(s)} = 0$.

7. Полагаем $k = k + 1$ и переходим к п. 5 до тех пор, пока $k \leq N$.

8. Полагаем $j = j + 1$ и переходим к п. 4 до тех пор, пока $j \leq m$.

9. Полагаем $i = i + 1$ и переходим к п. 3 до тех пор, пока $i \leq k$.

10. Вывод матрицы $A^{(s)} = [a_{ij}^{(s)}]$ и построенного на ее основе представления реализации отношения (34) в виде двудольного графа.

11. Полагаем $s = s + 1$ и переходим к п.2 до тех пор, пока $s \leq S$.

Введем в рассмотрение характеристики сгенерированных реализаций отношения (34).

Из теории графов известно, что число дуг, исходящих из некоторой i -й вершины графа вычисляется как:

$$d_i = \sum_{j=1}^m a_{ij}, \quad i = (\overline{1, n}), \quad (35)$$

а число дуг, входящих в j -ю вершину, определяется выражением вида:

$$d_j = \sum_{i=1}^n a_{ij}, \quad j = (\overline{1, m}). \quad (36)$$

В этих формулах величинами a_{ij} являются элементы матрицы смежности рассматриваемого графа.

Используя выражения (35), (36), для каждой s -й реализации отношения \mathcal{R} можно определить номера элементов множества X , из которых «исходит» соответственно минимальное и максимальное число дуг орграфа.

Номера этих вершин определяются как:

$$i_{\min} = \min_{1 \leq i \leq n} d_i; \quad i_{\max} = \max_{1 \leq i \leq n} d_i. \quad (37)$$

Номера элементов множества Y , в которые «входят» соответственно минимальное и максимальное число дуг, определяются выражениями вида:

$$j_{\min} = \min_{1 \leq j \leq m} d_j; \quad j_{\max} = \max_{1 \leq j \leq m} d_j. \quad (38)$$

При выполнении N генераций формируемого отношения \mathcal{R} получаем с помощью выражений (35), (36) совокупности чисел $d_i^{(s)}$ и $d_j^{(s)}$, $s = (\overline{1, N})$, $i = (\overline{1, n})$, $j = (\overline{1, m})$. Тогда среднее число связей элемента $x_i \in X$ при N реализациях отношения (34) вычисляется как:

$$\bar{d}_i = \frac{1}{N} \sum_{s=1}^N d_i^{(s)}. \quad (39)$$

Аналогично, среднее число взаимодействий элемента $y_j \in Y$ с элементами множества X определяется по формуле вида:

$$\bar{d}_j = \frac{1}{N} \sum_{s=1}^k d_j^{(s)}. \quad (40)$$

Отметим, что в различных приложениях введенного случайного отношения \mathcal{R} для задания значений p_i и q_j используются такие законы распределения как равномерное распределение, закон Зипфа, «бинарное» распределение, распределение Парето, геометрическое распределение, биномиальное распределение и др.

Самостоятельная работа.

Изучить смысл и области применения этих законов распределения.

Используя подход из работы [Цегелик Г.Г.], можно предложить общие формулы для задания вероятностей p_i и q_j в процессе генерации S реализаций отношения (34):

$$p_i^{(s)} = \frac{1}{s^{c_i} H(c_i, S)}, \quad q_j^{(s)} = \frac{1}{s^{c_j} H(c_j, S)}, \quad s = (\overline{1, S}),$$

где c_i и c_j лежат в интервале $[0, 2]$, а зависимости H имеют следующий вид:

$$H(c_i, S) = \sum_{r=1}^S \left(\frac{1}{r^{c_i}} \right); \quad H(c_j, S) = \sum_{r=1}^N \left(\frac{1}{r^{c_j}} \right).$$

В этой работе утверждается, что при $c_j = 0$ получаем равномерное распределение, при $c_j = 1$ – закон Зипфа, а при $c_j = 0,8614$ – закон Парето («80-20»). «Бинарное» распределение элементов множеств X и Y записывается как:

$$p_i^{(s)} = \frac{1}{2^s}, \quad s = (\overline{1, S-1}), \quad p_i^{(S)} = \frac{1}{2^{S-1}}, \quad i = (\overline{1, n});$$

$$q_j^{(s)} = \frac{1}{2^s}, \quad s = (\overline{1, S-1}), \quad q_j^{(S)} = \frac{1}{2^{S-1}}, \quad j = (\overline{1, m}).$$

Пример.

Пусть имеется некоторая группа $X = \{x_1, x_2, x_3\}$ работников, которая должна выполнять множество $Y = \{y_1, y_2, y_3, y_4\}$ работ.

Возможности (компетенции) работников по их выполнению заданы матрицей:

$$A = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} & \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \end{matrix}.$$

Предположим, что в силу определенных причин (болезнь, отпуск, командировка) работник не вышел на работу. Будем считать, что i -й работник с вероятностью p_i , $i = (\overline{1, 3})$ находится на рабочем месте. Пусть значения этих вероятностей равны:

$$p_1 = 0,25; \quad p_2 = 0,5; \quad p_3 = 0,25.$$

Матрица \mathcal{P} , описывающая рассматриваемое отношение \mathcal{R} будет иметь вид:

$$\mathcal{P} = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} & \begin{bmatrix} 0,05 & 0 & 0,05 & 0 \\ 0 & 0,20 & 0 & 0 \\ 0 & 0,10 & 0,05 & 0,05 \end{bmatrix} \end{matrix}.$$

Предположим, что задания на выполнение работ множества Y поступают каждый рабочий день с вероятностями:

$$q_1 = \frac{1}{5} = 0,2; \quad q_2 = \frac{2}{5} = 0,4; \quad q_3 = \frac{1}{5} = 0,2; \quad q_4 = \frac{1}{5} = 0,2.$$

Построим с использованием приведенного алгоритма первую ($s = 1$) реализацию отношения \mathcal{R} при $N = 4$.

Конкретизируем неравенства, входящие в п. 5 алгоритма:

$$i = 1: \quad 0 < \alpha_k^{(s)} < 0,25;$$

$$i = 2: \quad 0,25 < \alpha_k^{(s)} < 0,75;$$

$$i = 3: \quad 0,75 < \alpha_k^{(s)} < 1;$$

$$j = 1: \quad 0 < \beta_k^{(s)} < 0,2;$$

$$j = 2: \quad 0,2 < \beta_k^{(s)} < 0,6;$$

$$j = 3: \quad 0,6 < \beta_k^{(s)} < 0,8;$$

$$j = 4: \quad 0,8 < \beta_k^{(s)} < 1.$$

Пусть ДСЧ сгенерировал следующие случайные числа:

$$\alpha_1^{(1)} = 0,5296; \quad \alpha_2^{(1)} = 0,2771; \quad \alpha_3^{(1)} = 0,8499; \quad \alpha_4^{(1)} = 0,07226;$$

$$\beta_1^{(1)} = 0,3579; \quad \beta_2^{(1)} = 0,1637; \quad \beta_3^{(1)} = 0,7891; \quad \beta_4^{(1)} = 0,9154.$$

Проводя анализ выполнения приведенных неравенств, имеем следующие результаты:

$$k = 1: \quad (x_3, y_2) \in \mathcal{R}_1, \quad a_{32}^{(1)} = 1;$$

$$k = 2: \quad (x_2, y_1) \in \mathcal{R}_1, \quad a_{21}^{(1)} = 1;$$

$$k = 3: (x_3, y_3) \in \mathcal{R}_1, \quad a_{33}^{(1)} = 1;$$

$$k = 4: (x_1, y_4) \in \mathcal{R}_1, \quad a_{14}^{(1)} = 1.$$

Формально реализация \mathcal{R}_1 представляется матрицей вида:

$$A^{(1)} = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} & \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}.$$

Из сравнения матриц A и $A^{(1)}$ следует, что из поступивших 4 заданий может быть выполнено только задание y_3 работником x_3 .

Самостоятельная работа.

Привести пример случайного отношения из области информатики.

Предложенные в Лекциях №6 и №7 матричные методы представления отношений и операций над отношениями наряду с их графическими представлениями, а также выражениями (35)-(38) предлагается использовать при решении разнообразных задач компьютерного анализа информационных систем и процессов, описываемых их теоретико-множественными моделями.

Рассмотренные выше отношения предполагают, что они формируются исследователем при их матричном описании или из решения специальных задач булевого программирования.

Характерной особенностью таких отношений является тот факт, что они являются неизменными в процессе решения на их основе всевозможных задач анализа сложных информационных систем. Последнее означает, что, например, при использовании бинарного отношения (4) состав входящих в не-

го множеств X и Y , а также элементы матрицы A , определяемые выражением (5), не изменяются при решении таких задач.

Практика разработки и применения теоретико-множественных моделей в различных задачах потребовала дополнительно к рассмотренному введению в теорию отношений следующих неклассических видов отношений дискретных множеств:

- условные отношения;
- динамические отношения;
- стохастические отношения.

Далее в трех последующих разделах для бинарных отношений вида (4) будут рассмотрены определения и алгоритмы построения этих новых видов отношений.

9. Условные бинарные отношения

Использование бинарных отношений при решении различных прикладных задач показало, что состав множеств X и Y , а также значения элементов матрицы A могут изменяться в зависимости от определенных условий их решения. Такие отношения будем называть условными. В этой связи все отношения, рассмотренные в Лекции №5, будем считать безусловными отношениями.

Условное бинарное отношение, формируемое на дискретных множествах $X = \{x_i | i = \overline{1, k}\}$ и $Y = \{y_j | j = \overline{1, m}\}$, с использованием некоторого условия u записывается как:

$$R(u) \subseteq X_u \times Y_u, \quad (41)$$

где u – используемое условие; $X_u \subseteq X$; $Y_u \subseteq Y$ - параметрические множества.

В качестве условия, используемого при формировании отношения $R(u)$, выступают логические, геометрические,

экономические и другие требования, определяющие условия его существования. Отношение (41) будем описывать с помощью условной матрицы $A(u) = [a_{ij}(u)]_{k \times m}$, элементы которой принимают следующие значения:

$$a_{ij}(u) = \begin{cases} 1, & \text{если для упорядоченной пары } (x_i, y_j), \\ & \text{где } x_i \in X, y_j \in Y, \text{ выполняется условие «}u\text{»}; \\ 0 & \text{в противном случае.} \end{cases} \quad (42)$$

Если рассматриваемое условие не выполняется ни для одной пары (x_i, y_j) , где $x_i \in X, y_j \in Y$, т.е. матрица $A(u)$ является нулевой матрицей, то считается, что отношения (41) не существует и оно является пустым множеством ($R(u) = \emptyset$). С точки зрения представления бинарного отношения (4) в форме двудольного графа считается, что в этом случае граф такого отношения состоит из несвязанных между собой изолированных элементов множеств X и Y .

Рассмотрим пример условного бинарного отношения, возникшего при решении задач планирования процессов стрельбы управляемыми артиллерийскими снарядами с лазерной головкой самонаведения. При этом «подсвет» целей осуществлялся с беспилотного разведывательно-корректировочного вертолета [Козар А.Н., Моисеев В.С.].

Пример.

Пусть множество X включает в себя множество артиллерийских орудий первого (x_1) и второго (x_2) типов. Множество $Y = \{y_1, y_2, y_3, y_4\}$ описывает совокупность наименований разведанных целей, по которым должны нанести удар имеющиеся орудия.

Будем считать заданной матрицу $L = [l_{ij}]$ расстояний между каждым x_i -м орудием y_j -й целью, $i = (\overline{1, 2})$, $j = (\overline{1, 4})$.

Каждому орудью x_i поставим в соответствие величину d_i максимальной дальности его стрельбы, $i = (\overline{1,2})$.

Выражение (42) в этом случае конкретизируется как:

$$a_{ij}(u) = \begin{cases} 1, & \text{если для упорядоченной пары } (x_i, y_j), \\ & \text{где } x_i \in X, y_j \in Y, \text{ выполнено неравенство} \\ & l_{ij} \leq d_i; \\ 0, & \text{если для упорядоченной пары } (x_i, y_j), \\ & \text{где } x_i \in X, y_j \in Y, \text{ имеет место неравенство} \\ & l_{ij} > d_i. \end{cases} \quad (43)$$

В этом примере условие существования отношения между доступными целями и имеющимися орудиями с учетом дальности их стрельбы конкретизируется как:

$$u: l_{ij} \leq d_i, \quad i = (\overline{1,2}), \quad j = (\overline{1,4}). \quad (44)$$

Пусть по результатам обработки разведывательных данных сформирована следующая матрица взаимных расстояний между орудиями и обнаруженными целями:

$$L = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 5 & 25 & 13 & 10 \\ 22 & 14 & 16 & 9 \end{bmatrix} \end{matrix},$$

в которой значения элементов l_{ij} заданы в километрах. Рассматриваемые орудия x_1 и x_2 имеют следующие характеристики дальности стрельбы:

$$d_1 = 12 \text{ км}; \quad d_2 = 15 \text{ км}.$$

Тогда, перебирая всевозможные сочетания значений индексов $i \in (\overline{1,2})$, $j \in (\overline{1,4})$ и используя неравенство (44) как условие, получаем следующий вид матрицы, описывающей формируемое отношение «орудия-цели»:

$$A(u) = \begin{matrix} & y_1 & y_2 & y_3 & y_4 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 5 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \end{matrix}.$$

Графическое представление отношения $R(u)$, построенное на основе этой матрицы, приведено на Рис. 34.

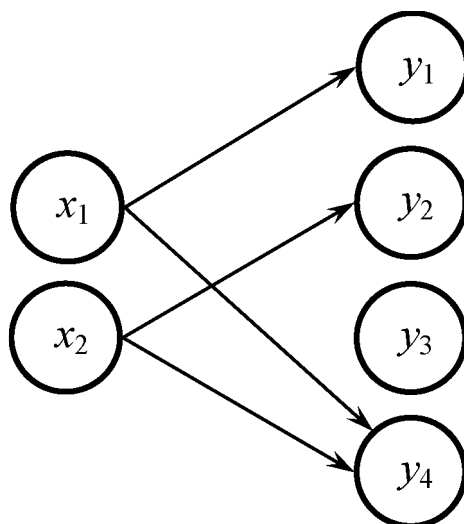


Рис. 34

Из этого рисунка следует, что орудие x_1 может вести огонь по целям y_1 и y_4 , а орудие x_2 – по целям y_2 и y_4 . По цели y_3 при имеющихся исходных данных вести огонь не может ни одно из имеющихся орудий.

Таким образом, множества, входящие в правую часть выражения (41) конкретизируются в рассматриваемом примере как:

$$X_u = X, Y_u = \{y_1, y_2, y_4\} \subset Y.$$

Отметим, при изменении значений матрицы L может измениться как состав множеств X_u, Y_u , так и значения элементов матрицы $A(u)$.

Самостоятельная работа.

Привести примеры условных отношений из области прикладной информатики.

На практике условие u может зависеть от ряда случайных параметров:

$$u = u(V, W, \dots),$$

где параметры V, W, \dots являются случайными событиями или случайными величинами с заданными вероятностными характеристиками (законами распределения).

Это позволяет говорить о случайных условных отношениях:

$$\mathcal{R}(u) \subseteq X_u \times Y_u.$$

В таких отношениях состав множеств X_u и Y_u в общем случае будет являться случайным.

Для применения в моделях процессов и систем отношений $\mathcal{R}(u)$ используется алгоритм, приведенный в Лекции №8.

Проиллюстрируем смысл введенного отношения на продолжении приведенного выше Примера.

Будем считать, что рассматриваемые орудия размещены на плоскости в точках с координатами (a_i, b_i) , $i = (\overline{1, 2})$.

Координаты маневрирующих целей описываются случайными векторами (V_j, W_j) с известными плотностями распределения $f_j(v, w)$, $j = (\overline{1, 4})$.

Отметим, что расстояния между орудиями и целями будут случайными величинами, определяемыми как:

$$\mathcal{L}_{ij} = \sqrt{(V_j - a_i)^2 + (W_j - b_i)^2}, \quad i = (\overline{1, 2}), \quad j = (\overline{1, 4}).$$

Тогда условия (44) примут вид:

$$u: \mathcal{L}_{ij} \leq d_i, \quad i = (\overline{1, 2}), \quad j = (\overline{1, 4}).$$

Для построения элементов матрицы $A(u) = [a_{ij}(u)]_{2 \times 4}$ в процессе генерации N реализаций отношения $\mathcal{R}(u)$ с помощью алгоритма из Лекции №8 необходимо определить значения вероятностей p_{ij} того, что j -я цель будет достижима при стрельбе из i -го орудия, $i = (\overline{1, 2}), j = (\overline{1, 4})$.

Расчетная схема для их вычисления приведена на Рис. 35.

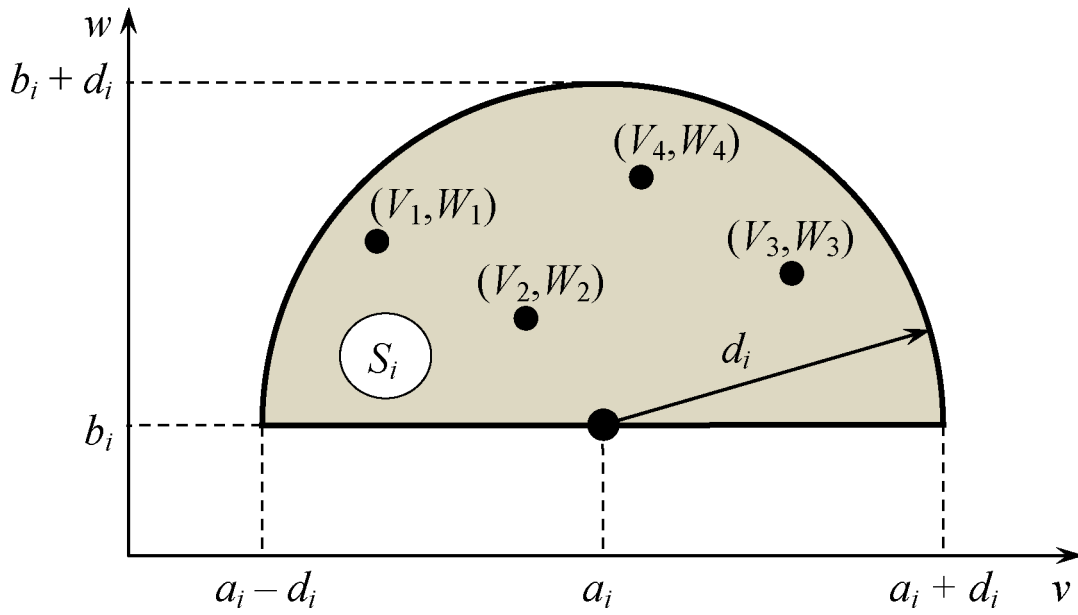


Рис. 35

На этом рисунке представлены координаты i -го орудия и случайных точек возможных мест расположения целей. Заштрихованная область S_i определяет область достижения целей i -м орудием при максимальной дальности стрельбы d_i , $i = (\overline{1, 2})$. Граница этой области задается выражением:

$$w = \begin{cases} b_i; \\ b_i + \sqrt{d_i^2 - (v_j - a_i)^2} \end{cases}$$

при $v \in [a_i - d_i; a_i + d_i]$.

При этом предполагается, что i -е орудие производит стрельбу только по целям, координаты которых удовлетворяют условиям:

$$V_j \geq a_i, \quad i = (\overline{1, 2}), \quad j = (\overline{1, 4}).$$

Из теории вероятностей известно, что вероятность попадания случайной точки в заданную плоскую область вычисляется как:

$$p_{ij} = P\{(V_j, W_j) \in S_i\} = \iint_{S_i} f_j(v, w) dv dw.$$

Входящий в эту формулу двойной интеграл конкретизируется с учетом приведенного выше уравнения границы области S_i следующим образом:

$$p_{ij} = \int_{a_i-d_i}^{a_i+d_i} dv \int_{b_i}^{b_i+\sqrt{d_i^2-(v-a_i)^2}} f_j(v,w)dw, \quad i = (\overline{1,2}), \quad j = (\overline{1,4}).$$

Самостоятельная работа.

Привести пример случайного условного отношения из области прикладной информатики.

10. Динамические бинарные отношения

Введем понятие динамического отношения, которое может рассматриваться как обобщение введенного понятия условного отношения.

Пусть задан некоторый интервал времени $[t_0, t_1]$, на котором предполагается рассмотрение формируемого динамического отношения.

Динамическое бинарное отношение будем представлять выражением вида:

$$R^t \subseteq X^t \times Y^t, \quad t \in [t_0, t_1]. \quad (45)$$

Здесь t - текущий момент времени; $X^t \subseteq X$, $Y^t \subseteq Y$ - параметрические множества, формируемые в момент времени t на основе заданных дискретных множеств $X = \{x_i | i = (\overline{1,k})\}$ и $Y = \{y_j | j = (\overline{1,m})\}$.

Для каждой упорядоченной пары элементов (x_i, y_j) этих множеств введем в рассмотрение используемую в математической логике предикатную функцию $\mathcal{P}_{ij}(t)$, $i = (\overline{1,k})$, $j = (\overline{1,m})$, принимающую в момент времени $t \in [t_0, t_1]$ следующие значения:

$$\mathcal{P}_{ij}(t) = \begin{cases} 1, \text{ если } (x_i, y_j) \subseteq R^t; \\ 0, \text{ если } (x_i, y_j) \notin R^t. \end{cases} \quad (46)$$

При решении практических задач предикатная функция $\mathcal{P}_{ij}(t)$ строится с использованием параметров и характеристик элементов $x_i \in X$ и $y_j \in Y$, определяющих физический, технический, экономический и другой смысл формируемого отношения (45).

Будем считать, что при формировании функции $\mathcal{P}_{ij}(t)$ используются условия вида:

$$\varphi_{ij}(c_i(t), b_j(t)) \leq 0, \quad i = (\overline{1, k}), \quad j = (\overline{1, m}), \quad (47)$$

где $c_i(t)$ и $b_j(t)$ – векторы параметров и характеристик элементов $x_i \in X$, $y_j \in Y$.

При этом предполагается, что если в момент времени t для определенных значений $i \in (\overline{1, k}), j \in (\overline{1, m})$ неравенства системы (47) выполняются, то предикатная функция $\mathcal{P}_{ij}(t) = 1$. Если хотя бы одно из этих неравенств в рассматриваемый момент времени t не выполняется, то $\mathcal{P}_{ij}(t) = 0$.

Следуя выражению (46), отношение (45) будем описывать матричной функцией $A(t) = [a_{ij}(t)]_{k \times m}$, элементы которой определяются как:

$$a_{ij}(t) = \mathcal{P}_{ij}(t), \quad i = (\overline{1, k}), \quad j = (\overline{1, m}). \quad (48)$$

Рассмотрим основные случаи динамического взаимодействия элементов множеств X и Y .

Пусть некоторые элементы $x_i \in X$, $y_j \in Y$ могут вступать во взаимодействие между собой только в один момент времени $t = \tau_{ij} \geq t_0$ и сохранять его до момента времени t_1 .

Для любой пары (x_i, y_j) этот момент времени определяется как:

$$\tau_{ij} = \arg\{\mathcal{P}_{ij}(t) = 1; i = (\overline{1, k}), j = (\overline{1, m}), t \in [t_0, t_1]\}. \quad (49)$$

В этом случае функция $\mathcal{P}_{ij}(t)$ примет следующие значения:

$$\mathcal{P}_{ij}(t) = \begin{cases} 1, \text{ при } t \in [t_0, \tau_{ij}]; \\ 0, \text{ при } t \notin [\tau_{ij}, t_1]. \end{cases} \quad (50)$$

Графическое представление соответствующих элементов матрицы $A(t)$ для этого случая, построенного с учетом выражений (48) и (49), приведено на Рис. 36.

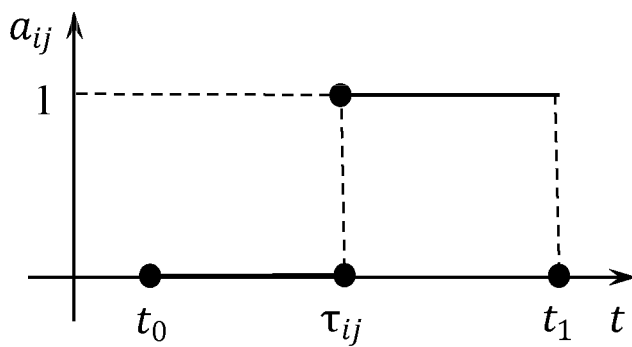


Рис. 36

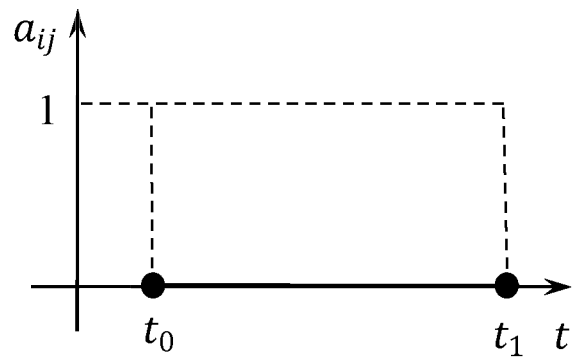


Рис. 37

Отметим, что функция $a_{ij}(t)$ в точке $t = \tau_{ij}$ имеет разрыв 1-го рода. Это означает, что $a_{ij}(\tau_{ij}) = 1$.

Если для некоторых $i \in (\overline{1, k})$ и $j \in (\overline{1, m})$ не существует решения уравнения $\mathcal{P}_{ij}(t) - 1 = 0$, то есть неравенства (47) не выполняются для всех значений $t \in [t_0, t_1]$, то имеем случай, когда функция $\mathcal{P}_{ij}(t) \equiv 0$ и соответствующие элементы матрицы $A(t)$ тождественно равны нулю.

Графическое представление таких элементов матрицы $A(t)$ приведено на Рис. 37.

Рассмотрим случай, когда определенные элементы $x_i \in X$, $y_j \in Y$ могут вступать на интервале времени $[t_0, t_1]$ во взаимодействие несколько раз.

Введем следующие обозначения:

$\tau_{ij}^{(n)}$ - время начала n -взаимодействия этих элементов;

$\theta_{ij}^{(n)}$ - время завершения их взаимодействия, где $n = 1, 2, 3, \dots$

Значения этих параметров будем определять из выражений вида:

$$\begin{aligned}\tau_{ij}^{(n)} &= \arg \left\{ \mathcal{P}_{ij}(t) = 1, \quad t \geq \theta_{ij}^{(n-1)}, \quad t \in [t_0, t_1] \right\}; \\ \theta_{ij}^{(n)} &= \arg \left\{ \mathcal{P}_{ij}(t) = 0, \quad t \geq \tau_{ij}^{(n-1)}, \quad t \in [t_0, t_1] \right\},\end{aligned}\quad (51)$$

где $n = 1, 2, 3, \dots$

При этом предполагается, что $\theta_{ij}^{(0)} = t_0$ и $\tau_{ij}^{(0)} = t_0$.

Предельное число N взаимодействий рассматриваемых элементов определяется выполнением следующих неравенств:

$$\tau_{ij}^{(N)} \leq t_1 \leq \theta_{ij}^{(N)}, i = (\overline{1, k}), j = (\overline{1, m}). \quad (52)$$

Взаимодействие элементов $x_i \in X, y_j \in Y$ описывается разрывными функциями вида:

$$a_{ij}(t) = \begin{cases} 1, & \text{при } t \in [\tau_{ij}^{(n)}, \theta_{ij}^{(n)}]; \\ 0, & \text{при } t \in [\theta_{ij}^{(n)}, \tau_{ij}^{(n+1)}], \end{cases} \quad (53)$$

где $n = 1, 2, 3, \dots, N$.

Графическое представление функций $a_{ij}(t)$, описываемых этим выражением, приведено на Рис. 38.

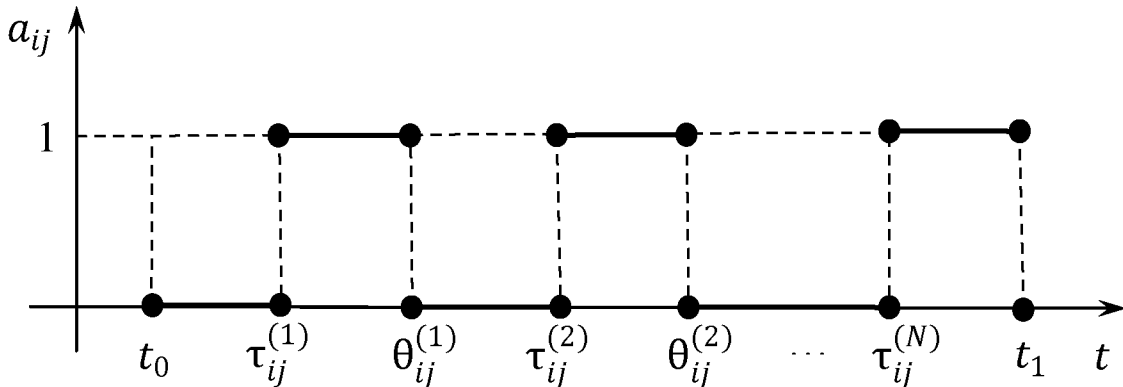


Рис. 38

Множества X^t и Y^t , входящие в отношение (45), формируются по известной матрице $A(t)$ следующим образом:

1. Если для всех моментов времени $t \in [t_0, t_1]$ все функции $a_{ij}(t) \equiv 1$, $i = (\overline{1, k})$, $j = (\overline{1, m})$, то имеем, что $X^t = X$ и $Y^t = Y$, $t \in [t_0, t_1]$. В этом случае отношение R^t называется статическим отношением и является классическим бинарным отношением вида (4).

2. Для значений функций $a_{ij}(t)$, представленных на Рис. 36-Рис. 38 эти множества для любого значения $t \in [t_0, t_1]$ формируются как:

$$\begin{aligned} X^t &= \{x_i \in X | a_{ij}(t) = 1, \quad i = (\overline{1, k}), \quad j = (\overline{1, m})\}; \\ Y^t &= \{y_j \in Y | a_{ij}(t) = 1, \quad i = (\overline{1, k}), \quad j = (\overline{1, m})\}. \end{aligned} \quad (54)$$

Процесс формирования динамического отношения (45) проиллюстрируем следующим примером.

Пример.

Пусть имеются две группы находящихся в воздухе самолетов, типы которых задаются множествами $X = \{x_1, x_2\}$ и $Y = \{y_1, y_2, y_3\}$.

Каждый из самолетов $x_i \in X$ и $y_j \in Y$ описывается следующими векторами их характеристик:

$$\begin{aligned} c_i(t) &= (\varepsilon_i(t), \eta_i(t), \xi_i(t), d_i), i = (\overline{1, 2}); \\ b_j(t) &= (\varepsilon_j(t), \eta_j(t), \xi_j(t)), j = (\overline{1, 3}). \end{aligned} \quad (55)$$

Здесь $\varepsilon_i(t), \eta_i(t), \xi_i(t)$ и $\varepsilon_j(t), \eta_j(t), \xi_j(t)$ - пространственные декартовы координаты i -го и j -го самолетов, описывающие их положение в пространстве в момент времени $t \in [t_0, t_1]$; d_i - дальность действия средств обнаружения и сопровождения воздушных целей, установленных на i -м самолете, $i = (\overline{1, 2})$.

Будем считать, что самолеты множеств X и Y движутся произвольным образом в некоторой области воздушного пространства. При этом задачей самолетов, входящих во множество X , является обнаружение и наблюдение за самолетами множества Y .

Требуется построить отношение вида (45), описывающее в каждый момент времени $t \in [t_0, t_1]$ состояние процесса обнаружения и сопровождения самолетов множества Y самолетами, входящими во множество X .

Построим предикатные функции $\mathcal{P}_{ij}(t)$, описывающие обнаружение и сопровождение i -м самолетом j -го самолета в момент времени $t \in [t_0, t_1]$.

Взаимное расстояние между этими самолетами в любой момент времени t определяется по формуле вида:

$$r_{ij} = \sqrt{(\varepsilon_i(t) - \varepsilon_j(t))^2 + (\eta_i(t) - \eta_j(t))^2 + (\xi_i(t) - \xi_j(t))^2}, \quad (56)$$

$$i = (\overline{1, 2}), \quad j = (\overline{1, 3}).$$

Неравенства (47) в нашем случае описывают условия того, что i -й самолет обнаружил и сопровождает в текущий момент времени j -ю цель.

Эти неравенства, с учетом выражения (56) для $i = 1$ и $j = 2$ конкретизируются как:

$$r_{1j}(t) \leq d_1; \quad r_{2j}(t) \leq d_2, \quad j = (\overline{1, 3}). \quad (57)$$

Предполагается, что если при некотором $t \in [t_0, t_1]$ оба эти неравенства выполняются для некоторого $j \in (\overline{1, 3})$, то это означает, что самолеты x_1 и x_2 обнаружили и сопровождают в этот момент времени некоторую цель $y_j \in Y$. Предикатные функции $\mathcal{P}_{ij}(t)$ в этом случае будут иметь следующие значения:

1) $\mathcal{P}_{1j}(t) = 1, \mathcal{P}_{2j}(t) = 1$ - для конкретного значения $j \in (\overline{1,3})$;

2) $\mathcal{P}_{1j}(t) = 0, \mathcal{P}_{2j}(t) = 0$ - для всех других значений $j \in (\overline{1,3})$.

Если в момент времени t для некоторого $j \in (\overline{1,3})$ выполняется только первое из неравенств (57), то $\mathcal{P}_{1j}(t) = 1, \mathcal{P}_{2j}(t) = 0$. В случае выполнения только второго неравенства из состава (57) получаем, что $\mathcal{P}_{1j}(t) = 0, \mathcal{P}_{2j}(t) = 1$.

При выполнении в момент времени t всех неравенств системы (57) имеем, что $\mathcal{P}_{ij}(t) = 1, i = (\overline{1,2}), j = (\overline{1,3})$.

Будем считать, что процесс обнаружения и сопровождения целей рассматривается на интервале времени $[t_0, t_1]$, где $t_0 = 5$ с, $t_1 = 100$ с.

Графическое представление функций $r_{ij}(t)$ на этом интервале времени, определяемых выражением (56), а также факты выполнения или невыполнения условий (57) в конкретные моменты времени $t \in [5, 100]$ с. приведены на Рис. 39.

Используя эти графики, отметим характерные особенности рассматриваемого процесса взаимодействия элементов множеств X и Y .

Из Рис. 39,а следует, что самолет x_1 обнаруживает цель y_1 при $t = 10$ с. и сопровождает ее до момента времени $t = 40$ с. Далее в интервале времени $[40, 60)$ с. эта цель выходит из его зоны обнаружения и сопровождения, в которую она возвращается на интервале времени $[60, 80)$ с. и на интервале $[80, 100]$ с. вновь покидает эту зону.

По Рис. 39,б и Рис. 39,д можно сделать вывод, что цель y_2 на интервале времени $[5, 100]$ с. не обнаруживается ни одним из самолетов x_1 и x_2 .

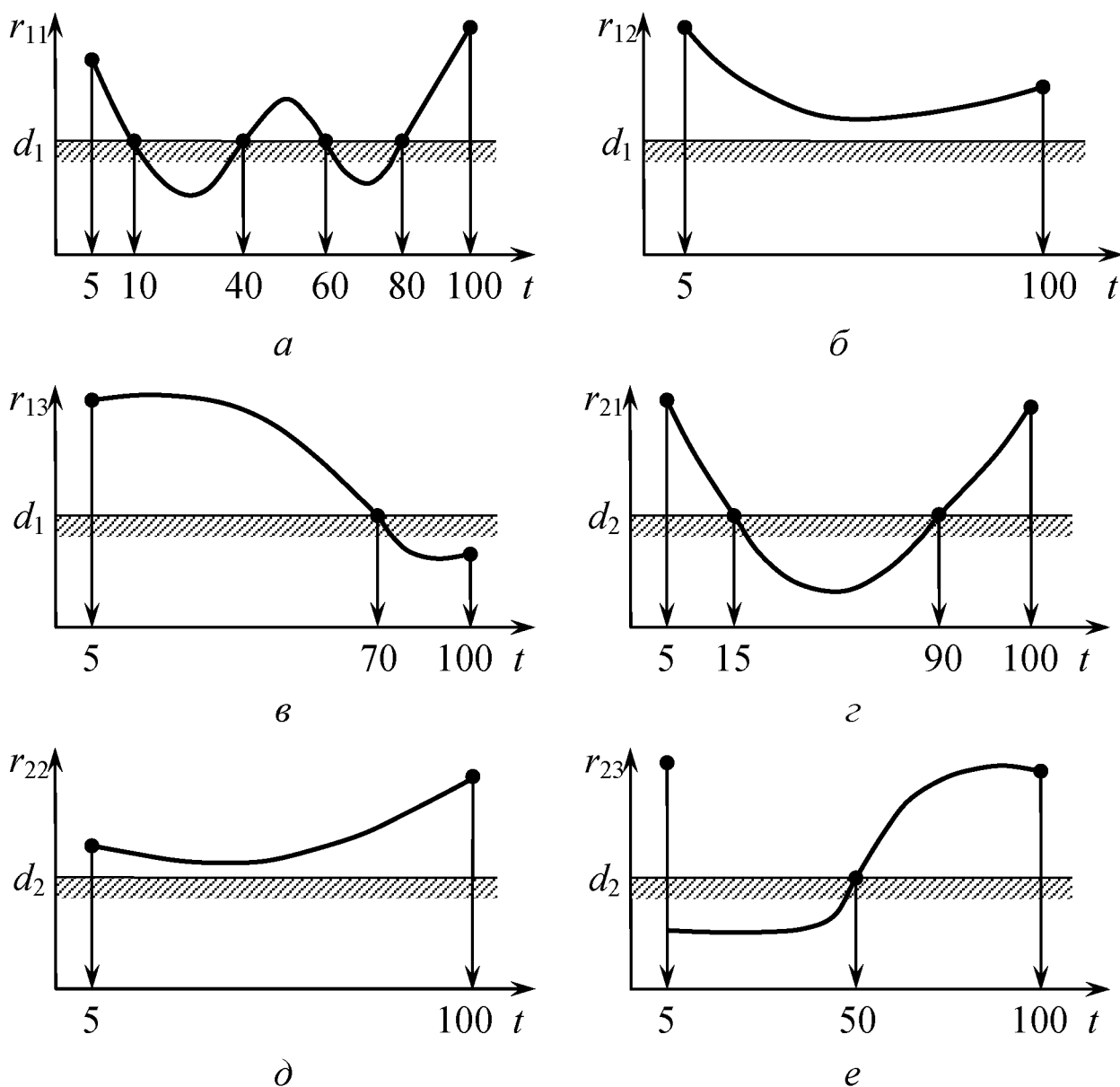


Рис. 39

Цель y_3 обнаруживается и сопровождается самолетом x_1 на интервале времени $[70, 100]$ с. (см. Рис. 39,в). Из Рис. 39,г и Рис. 39,е следует, что самолет x_2 обнаруживает и сопровождает цель y_1 на интервале времени $[15, 90)$ с., а цель y_3 – на интервале времени $[5, 50)$ с.

На Рис. 40 представлены графики элементов $a_{ij}(t)$ матричной функции $A(t)$, построенные на основе выражения (48) и приведенного на Рис. 39 представления процессов обнаружения и сопровождения целей множества Y .

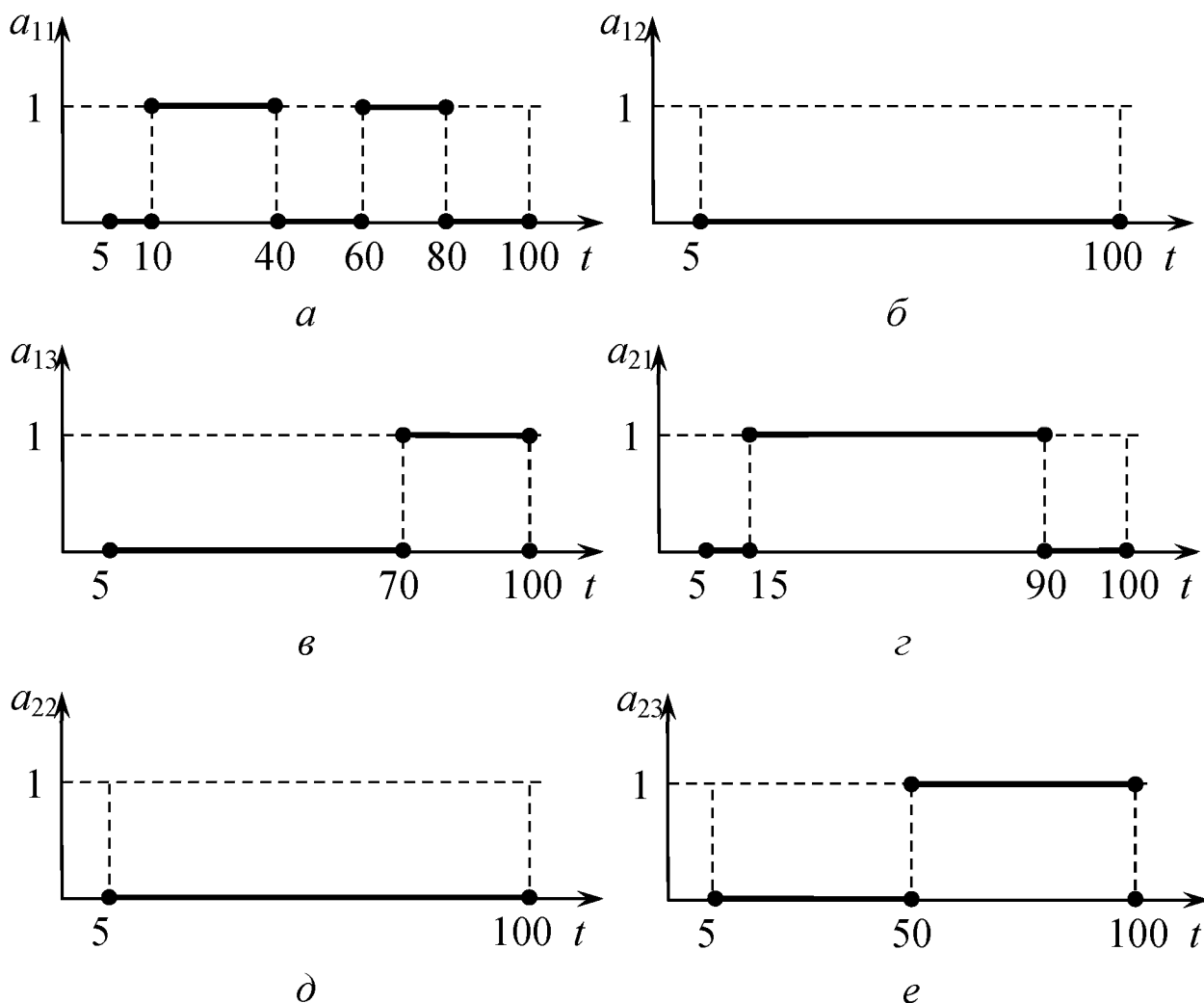


Рис. 40

На основе этих графиков для временной сетки с шагом $\Delta t = 5$ с построим состав параметрических множеств X^t и Y^t , определяемых матричной функцией $A(t)$ и выражениями (54), а также множество пар элементов $x_i \in X^t$ и $y_j \in Y^t$, составляющих формируемое отношение (45) (Табл. 1).

Из анализа этой таблицы следует, что на определенных промежутках времени, входящих в интервал $[5, 100]$ с. состав множеств X^t и Y^t , а так же пар (x_i, y_j) , не изменяется.

На Рис. 41 приведено графическое представление динамики состава отношения R^t на этих промежутках времени. Из этого рисунка следует, что на интервале $[5, 100]$ с. самолет $y_2 \in Y$ не обнаруживается самолетами множества X .

Таблица 1.

t, c	$a_{11}(t)$	$a_{12}(t)$	$a_{13}(t)$	$a_{21}(t)$	$a_{22}(t)$	$a_{23}(t)$	X^t	Y^t	$R^t \subseteq X^t \times Y^t$
5	0	0	0	0	0	1	x_2	y_3	(x_2, y_3)
10	1	0	0	0	0	1	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_2, y_3)$
15	1	0	0	1	0	1	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_2, y_1), (x_2, y_3)$
20	1	0	0	1	0	1	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_2, y_1), (x_2, y_3)$
25	1	0	0	1	0	1	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_2, y_1), (x_2, y_3)$
30	1	0	0	1	0	1	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_2, y_1), (x_2, y_3)$
35	1	0	0	1	0	1	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_2, y_1), (x_2, y_3)$
40	0	0	0	1	0	1	x_2	y_1, y_3	$(x_2, y_1), (x_2, y_3)$
45	0	0	0	1	0	1	x_2	y_1, y_3	$(x_2, y_1), (x_2, y_3)$
50	0	0	0	1	0	0	x_2	y_1	(x_2, y_1)
55	0	0	0	1	0	0	x_2	y_1	(x_2, y_1)
60	1	0	0	1	0	0	x_1, x_2	y_1	$(x_1, y_1), (x_2, y_1)$
65	1	0	0	1	0	0	x_1, x_2	y_1	$(x_1, y_1), (x_2, y_1)$
70	1	0	1	1	0	0	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_1, y_3), (x_2, y_1)$
75	1	0	1	1	0	0	x_1, x_2	y_1, y_3	$(x_1, y_1), (x_1, y_3), (x_2, y_1)$
80	0	0	1	1	0	0	x_1, x_2	y_1, y_3	$(x_1, y_3), (x_2, y_1)$
85	0	0	1	1	0	0	x_1, x_2	y_1, y_3	$(x_1, y_3), (x_2, y_1)$
90	0	0	1	0	0	0	x_1	y_3	(x_1, y_3)
95	0	0	1	0	0	0	x_1	y_3	(x_1, y_3)
100	0	0	0	0	0	0	x_1	y_3	(x_1, y_3)

Приведем общий алгоритм формирования отношения (45), который включает в себя следующие этапы:

1. Ввод значений моментов времени t_0 и t_1 , определяющих область существования отношения R^t ;

2. Ввод значения Δt для построения в интервале $[t_0, t_1]$ сетки значений моментов времени t в которых будет производиться формирование требуемого динамического отношения.

3. Полагаем $t = t_0$;
4. Полагаем $i = 1$;
5. Полагаем $j = 1$;
6. Проверка выполнения условий (47) с формированием индексного множества $J^t = \{(i, j) | i \in (\overline{1, k}), j \in (\overline{1, m})\}$, включающего в себя только те значения номеров элементов $x_i \in X$, $y \in Y$, для которых выполнены соответствующие неравенства из их состава;

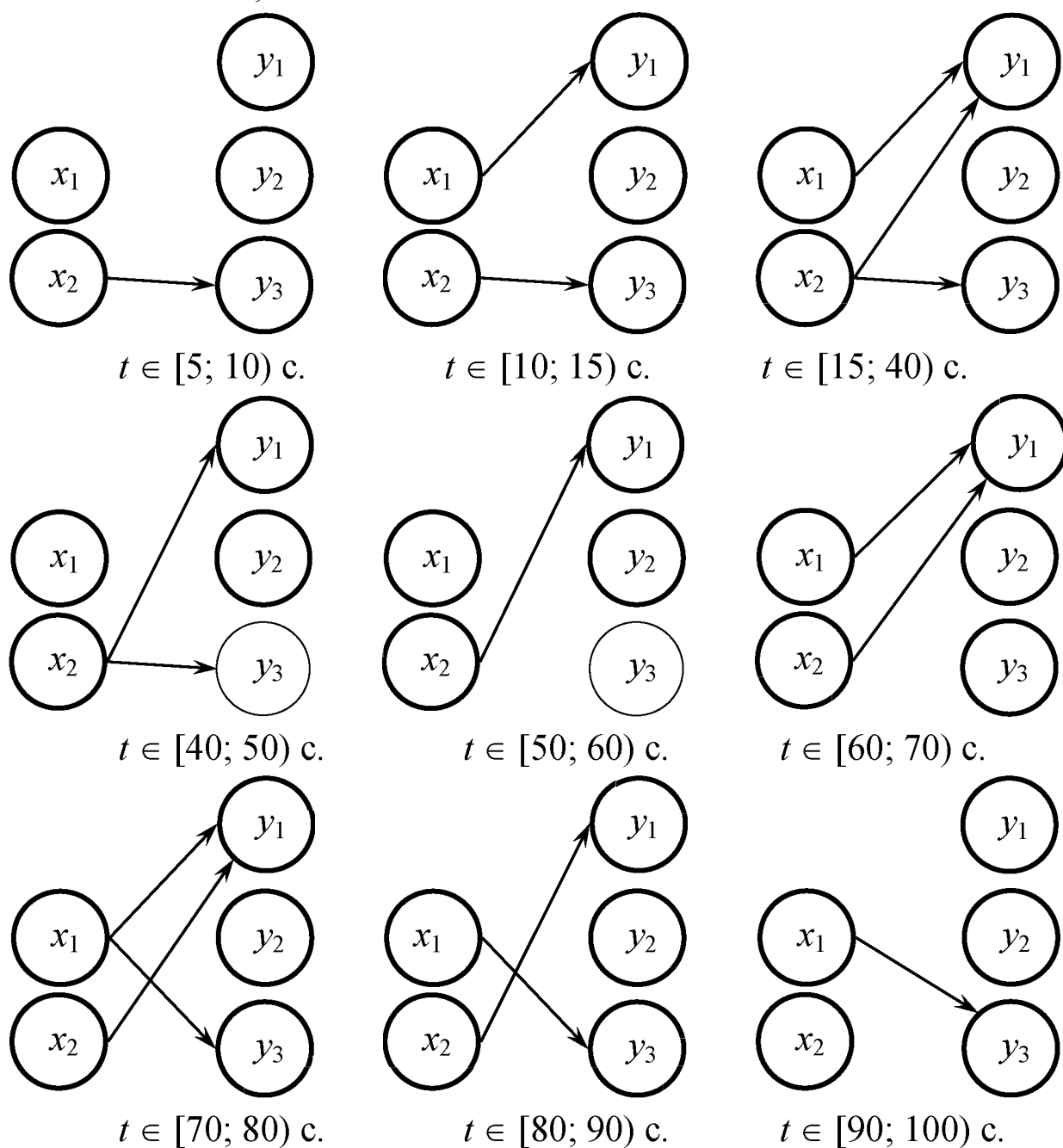


Рис. 41

7. Формирование элементов матрицы $A(t)$ по следующему правилу:

$$a_{ij}(t) = \begin{cases} 1, & \text{для пары } (i, j) \in J^t, \\ 0, & \text{для пары } (i, j) \text{ такой, что } J^t = \emptyset. \end{cases}$$

8. Полагаем $j = j + 1$ и переходим к п. 6 до тех пор, пока $j \leq m$;

9. Полагаем $i = i + 1$ и переходим к п. 5 до тех пор, пока $i \leq k$;

10. Полагаем $t = t + \Delta t$ и переходим к п. 4 до тех пор, пока $t \leq t_1$;

11. Используя сформированные значения матрицы $A(t)$, выражение (49) и индексные множества J^t , где $t \in [t_0, t_1]$, выдаем описание отношения R^t в виде Табл. 1 и в графическом виде (см. Рис. 41).

Самостоятельная работа

Привести пример динамического отношения из области прикладной информатики.

11. Стохастические бинарные отношения

Рассмотрим классическое бинарное отношение (4) множеств $X = \{x_1, x_2, \dots, x_k\}$ и $Y = \{y_1, y_2, \dots, y_m\}$.

Практика рассмотрения различных прикладных задач показала, что такие элементы могут вступать во взаимодействие с определенными вероятностями. Для описания таких взаимодействий предлагается использовать новый вид отношений, который будем называть стохастическими.

Будем считать, что для каждой упорядоченной пары (x_i, y_j) , $x_i \in X$, $y_j \in Y$ декартова произведения (4) заданы вероятности $p_{ij} \in [0, 1]$ их взаимодействия.

Для всех элементов рассматриваемых множеств эти вероятности могут быть сведены в матрицу вида:

$$P = \begin{matrix} & y_1 & y_2 & \dots & y_m \\ \begin{matrix} x_1 \\ x_2 \\ \dots \\ x_k \end{matrix} & \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \dots & \dots & \dots & \dots \\ p_{k1} & p_{k2} & \dots & p_{km} \end{bmatrix} \end{matrix}. \quad (58)$$

Если считать, что между каждым элементом $x_i \in X$ должна существовать связь с хотя бы одним элементом множества Y , то введенные вероятности должны удовлетворять следующим условиям:

$$\sum_{j=1}^m p_{ij} = 1, \quad i = (\overline{1, k}). \quad (59)$$

В этом случае, матрица P называется стохастической матрицей [Кофман А., Крюон Р.].

Стохастическое бинарное отношение, которое строится на основе матрицы P , будем обозначать следующим образом:

$$R^C \subseteq X^C \times Y^C, \quad (60)$$

где $X^C \subseteq X$; $Y^C \subseteq Y$ – случайные множества, состоящие из элементов исходных множеств X и Y .

Отметим, что отличие этого вида отношений от случайного отношения \mathcal{R} , введенного в Лекции №8, состоит в том, что в отношении R^C вероятности связей p_{ij} считаются заданными, а в отношении \mathcal{R} они вычисляются без требования выполнения условия (59).

На практике отношение R^C конкретизируется так же как и случайное отношение \mathcal{R} в виде совокупности реализаций:

$$R^C = \{R_1^C, R_2^C, \dots, R_s^C, \dots\}, \quad (61)$$

где R_s^C – s -я реализация отношения, которая получается в результате выполнения s -го опыта (эксперимента, испытания и т.п.).

Отметим, что каждая s -я реализация, входящая в выражение (61), является детерминированным отношением вида (4) и описывается соответствующей матрицей $A^{(s)} = [a_{ij}^{(s)}]_{k \times m}$, элементы которой для отношения R_S^C определяются выражением (5), где $s = 1, 2, 3, \dots$

Соответственно для каждой s -й реализации отношения (60) реализации множеств X^C и Y^C определяются как

$$\begin{aligned} X_s^C &= X \setminus \{x_i \in X \mid a_{ij}^{(s)} = 0, i \in (\overline{1, k}), j \in (\overline{1, m})\}; \\ Y_s^C &= Y \setminus \{y_j \in Y \mid a_{ij}^{(s)} = 0, i \in (\overline{1, k}), j \in (\overline{1, m})\}, \quad (62) \\ s &= 1, 2, 3, \dots \end{aligned}$$

Смысл этих выражений состоит в том, что из базовых множеств X и Y исключаются те элементы, которым в матрице $A^{(s)}$ соответствуют нулевые строки и нулевые столбцы.

Будем считать, что в процессе построения реализаций отношения (60) получено S его реализаций вида (61).

Используя построенные матрицы $A^{(s)}, s = (\overline{1, S})$, сформируем путем их сложения матрицу:

$$W = \sum_{s=1}^S A^{(s)}, \quad (63)$$

элементы которой вычисляются по формуле вида:

$$w_{ij} = \sum_{s=1}^S a_{ij}^{(s)}; \quad i = (\overline{1, k}), \quad j = (\overline{1, m}). \quad (64)$$

Исходя из смысла элементов матрицы A , можно утверждать, что величины w_{ij} определяют число упорядоченных пар (x_i, y_j) , $x_i \in X$, $y_j \in Y$, входящих в S реализаций отношения R^C , т.е. число взаимодействий этих элементов в S опытах.

Анализ значений этих величин позволяет определить пару элементов из множеств X и Y с номерами $i^* \in (\overline{1, k})$ и $j^* \in (\overline{1, m})$, которая имеет максимальное число взаимодействий в проведенных опытах. Такие номера могут быть определены из соотношения вида:

$$(i^*, j^*) = \arg \max_{\substack{1 \leq i \leq k, \\ 1 \leq j \leq m}} \{w_{ij}\}. \quad (65)$$

Оценки математических ожиданий числа элементов множеств X^C и Y^C , которые обозначим как \bar{N}_X^C и \bar{N}_Y^C , полученные по результатам S опытов, будем определять по следующим формулам:

$$\bar{N}_X^C = \frac{1}{S} \sum_{s=1}^S N_X^{(s)}, \quad \bar{N}_Y^C = \frac{1}{S} \sum_{s=1}^S N_Y^{(s)}, \quad (66)$$

где N_X^C и N_Y^C - мощности множеств X_S^C и Y_S^C , определяемые выражениями (62).

Оценки дисперсий случайных величин N_X^C и N_Y^C вычисляются как:

$$D[N_X^C] = \frac{1}{S-1} \sum_{s=1}^S (N_X^C - \bar{N}_X^C)^2, \quad (67)$$

$$D[N_Y^C] = \frac{1}{S-1} \sum_{s=1}^S (N_Y^C - \bar{N}_Y^C)^2.$$

Рассмотрим метод построения реализаций отношения R^C с использованием матрицы вероятностей (58).

Будем считать, что взаимодействие элементов из множеств X и Y происходит при наступлении совокупности определенных случайных событий.

Обозначим через B_{ij} случайное событие, вызывающее взаимодействие элемента $x_i \in X$ и элемента $y_j \in Y$. При наступлении этого события предполагается, что $(x_i, y_j) \in R^C$.

Вероятности наступления всех событий B_{ij} определяются как введенные выше вероятности p_{ij} , $i = (\overline{1, k})$, $j = (\overline{1, m})$.

Принимая во внимание условия (59), можно утверждать, что для каждого значения $i \in (\overline{1, k})$ и всех $j \in (\overline{1, m})$ события $B_{i1}, B_{i2}, \dots, B_{im}$ составляют полную группу случайных событий. Это означает, что в результате любого опыта одно из событий этой группы обязательно должно произойти.

Последнее позволяет использовать широко известный метод моделирования наступления случайных событий, который используется при статистическом моделировании процессов и систем [Вентцель Е.С.].

Для каждого из случайных событий $B_{i1}, B_{i2}, \dots, B_{im}$ построим отрезки на интервале $[0, 1]$ в форме, представленной на Рис. 42.

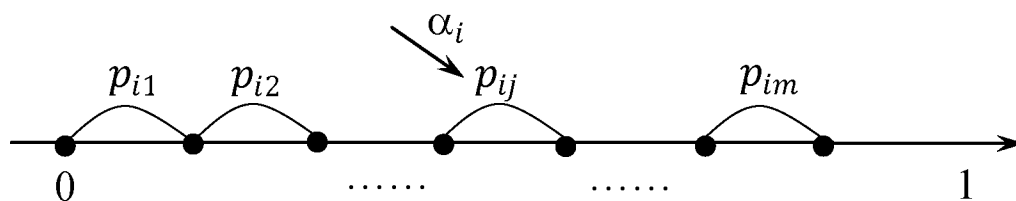


Рис. 42

Здесь длина каждого отрезка равна вероятности p_{ij} , $i = (\overline{1, k})$, $j = (\overline{1, m})$. Формально совокупности этих отрезков задаются множеством интервалов $(p_{ij}, p_{ij} + p_{ij+1})$, где $i = (\overline{1, k})$, $j = (\overline{0, m-1})$. При этом предполагается, что $p_{ij} = 0$ при $j = 0$ и $p_{ij} = 1$ при $j = m - 1$.

Из отмеченной выше работы следует, что если реализация равномерно распределенного случайного числа $\alpha_i \in (0, 1)$ по-

падает в интервал длиной p_{ij} (см. Рис. 42), то наступает случайное событие B_{ij} , $j = (\overline{1, m})$.

В этом случае считается, что в проведенном i -м опыте для пары $(x_i, y_j) \in R_s^C$ соответствующий элемент матрицы $A^{(s)}$ полагается равным единице.

Таким образом, общий алгоритм формирования S реализаций стохастического отношения (60) и вычисления его характеристик может быть описан следующей совокупностью этапов:

1. Вводим значения p_{ij} , $i = (\overline{1, k})$, $j = (\overline{1, m})$ матрицы вероятностей (58).

2. Задаём число S требуемых реализаций формируемого отношения R^C .

3. Полагаем $s = 1$.

4. Полагаем $i = 1$.

5. Генерируем с помощью датчика равномерно распределенных в интервале $(0,1)$ случайных чисел (ДСЧ) значение числа $\alpha_i^{(s)}$.

6. Полагаем $j = 0$.

7. Проверяем выполнение условия:

$$p_{ij} < \alpha_i^{(s)} < p_{ij} + p_{ij+1}.$$

8. Если это условие выполняется, полагаем $a_{i,j+1}^{(s)} = 1$, в противном случае $a_{i,j+1}^{(s)} = 0$.

9. Полагаем $j = j + 1$ и переходим к п. 7 до тех пор, пока $j \leq m - 1$.

10. Полагаем $i = i + 1$ и переходим к п. 5 до тех пор, пока $i \leq k$.

11. Формируем матрицу $A^{(s)} = [a_{ij}^{(s)}]_{k \times m}$.

12. Формируем множества X_s^C и Y_s^C согласно выражениям (62).

13. Полагаем $s = s + 1$ и переходим к п. 4 до тех пор, пока $s \leq S$.

14. Вычисляем с помощью выражений (63) и (64) матрицу W .

15. Вычисляем числовые характеристики (65)-(67).

16. Формируем результаты в табличной и графической формах.

Отметим, что данный алгоритм может использоваться для моделирования поведения стохастического отношения R^C на интервале времени $[t_0, t_1]$. В этом случае число реализаций S может быть вычислено по формуле:

$$S = \frac{t_1 - t_0}{\Delta t}, \quad (68)$$

где Δt - требуемый шаг моделирования.

В развитии введенного понятия стохастического отношения R^C можно рассмотреть случай, когда элементы матрицы (58) зависят от времени, т.е. формирование этого отношения осуществляется на основе матричной функции $P(t) = [p_{ij}(t)]_{k \times m}$, $t \in [t_0, t_1]$. В этом случае отношение (60) может рассматриваться как стохастическое динамическое отношение.

Пример.

Пусть заданы множество пользователей $X = \{x_1, x_2\}$ и множество серверов $Y = \{y_1, y_2, y_3\}$, с которыми они взаимодействуют в процессе своей работы. В составе множества Y могут рассматриваться, например, коммуникационный сервер (y_1), сервер баз данных (y_2) и сервер приложений (y_3).

Будем считать, что в результате сбора и обработки соответствующих статистических данных сформирована матрица

вида (58), описывающая относительные частоты (оценки вероятности) обращения пользователей к соответствующим серверам:

$$P = \begin{matrix} & y_1 & y_2 & y_3 \\ \begin{matrix} x_1 \\ x_2 \end{matrix} & \begin{bmatrix} 0,1 & 0,6 & 0,3 \\ 0,4 & 0,5 & 0,1 \end{bmatrix} \end{matrix}.$$

Построим, следуя приведенному алгоритму, $S = 5$ реализаций отношения (60), соответствующие дням рабочей недели пользователей x_1 и x_2 .

Пусть $s = 1$ и с помощью ДСЧ сгенерированы два числа $\alpha_1^{(s)} = 0,41$ и $\alpha_2^{(s)} = 0,67$.

Для пользователя $x_1 \in X$ связь с серверами системы определяется случайными событиями B_{11}, B_{12}, B_{13} . Из Рис. 43,а следует, что $\alpha_1^{(1)} \in (0,1; 0,7)$. Это означает, что реализовано событие B_{12} , т.е. его взаимодействие с сервером $y_2 \in Y$. Тогда элементы первой строки матрицы $A^{(1)}$ конкретизируются как:

$$a_{11}^{(1)} = 0; \quad a_{12}^{(2)} = 1; \quad a_{13}^{(3)} = 0.$$

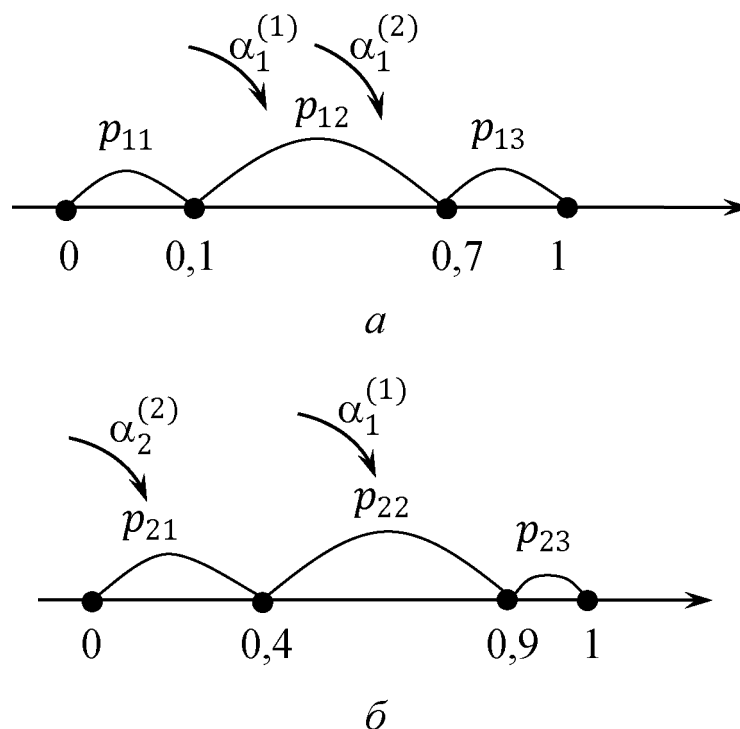


Рис. 43

Для пользователя $x_2 \in X$ из Рис. 43,б следует, что $\alpha_2^{(1)} \in (0,4; 0,9)$. Это означает, что произошло его взаимодействие с сервером $y_2 \in Y$ (событие B_{22}). В этом случае вторая строка матрицы $A^{(1)}$ записывается в виде:

$$a_{21}^{(1)} = 0; \quad a_{22}^{(2)} = 1; \quad a_{23}^{(3)} = 0.$$

На Рис. 44 приведены графическое представление отношения R_1^C и соответствующая ему матрица $A^{(1)}$.

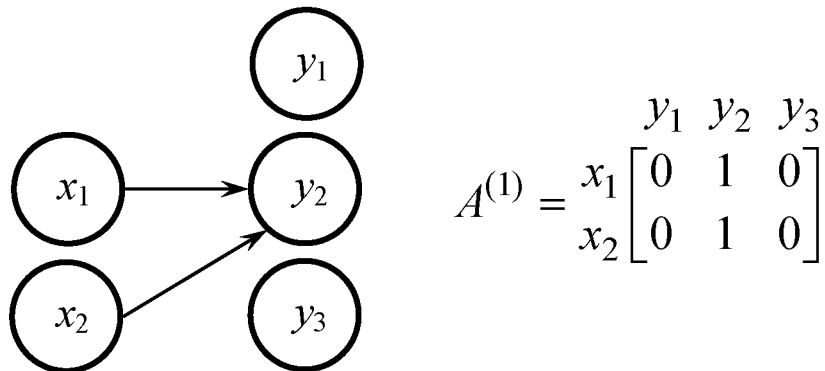


Рис. 44

Используя выражения (62), получаем следующий состав первой реализации множеств X^C и Y^C :

$$X_1^C = \{x_1, x_2\}; \quad Y_1^C = \{y_2\}.$$

Отсюда следует, что $N_X^{(1)} = 2$, $N_Y^{(1)} = 1$.

Для $s = 2$ и случайных чисел $\alpha_1^{(2)} = 0,34$; $\alpha_2^{(2)} = 0$ из Рис. 43 следует, что имели место случайные события B_{12} и B_{21} . Реализация R_2^C формируемого отношения представлена на Рис. 45.

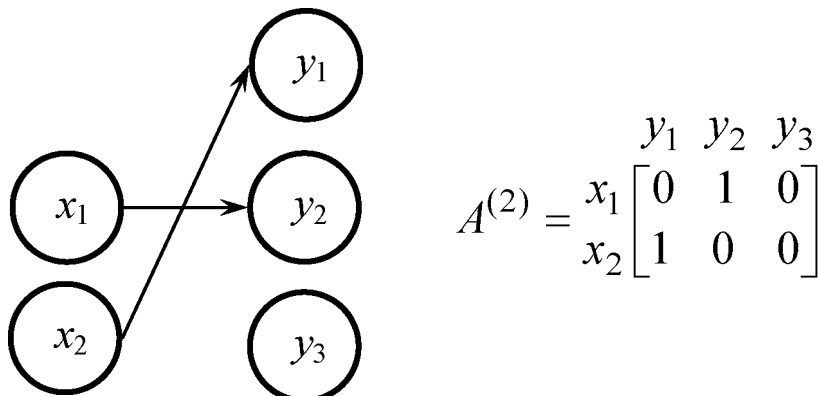


Рис. 45

Для этой реализации имеем, что $X_1^C = \{x_1, x_2\}$, $Y_1^C = \{y_1, y_2\}$ и $N_X^{(2)} = 2$, $N_Y^{(2)} = 2$.

На Рис. 46 приведены матричные и графические представления реализаций отношения R_s^C для $s = 3, 4, 5$ и использованные при этом значения случайных чисел $\alpha_1^{(s)}$ и $\alpha_2^{(s)}$.

Из сопоставления Рис. 45 и Рис. 46 следует, что реализации R_2^C и R_3^C совпадают между собой. Это означает, что во 2-й и 3-й дни рабочей недели пользователи взаимодействуют только с серверами y_1 и y_2 .

В Табл. 2 приведено описание полученных реализаций формируемого стохастического отношения.

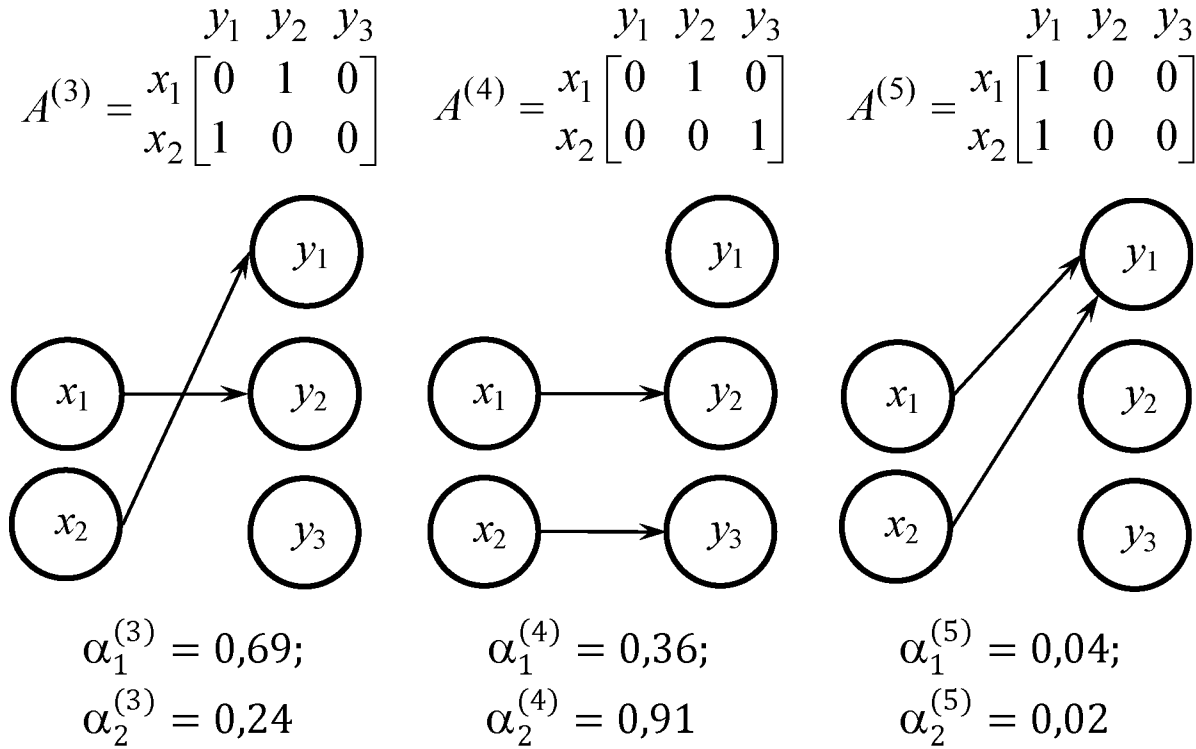


Рис. 46

Таблица 2

s	$a_{11}^{(s)}$	$a_{12}^{(s)}$	$a_{13}^{(s)}$	$a_{21}^{(s)}$	$a_{22}^{(s)}$	$a_{23}^{(s)}$	X_s^C	Y_s^C	$N_X^{(s)}$	$N_Y^{(s)}$	R_s^C
1	0	1	0	0	1	0	x_1, x_2	y_2	2	1	$(x_1, y_2), (x_2, y_2)$
2	0	1	0	1	0	0	x_1, x_2	y_1, y_2	2	2	$(x_1, y_2), (x_2, y_1)$
3	0	1	0	1	0	0	x_1, x_2	y_1, y_2	2	2	$(x_1, y_2), (x_2, y_1)$
4	0	1	0	0	0	1	x_1, x_2	y_2, y_3	2	2	$(x_1, y_2), (x_2, y_3)$
5	1	0	0	1	0	0	x_1, x_2	y_1	2	1	$(x_1, y_1), (x_2, y_1)$

Матрица W согласно выражениям (63), (64) определяется как:

$$\begin{aligned} W &= A^{(1)} + A^{(2)} + \dots + A^{(5)} = \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} + \\ &+ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{matrix} & y_1 & y_2 & y_3 \\ x_1 & 1 & 4 & 0 \\ x_2 & 3 & 1 & 1 \end{matrix}. \end{aligned}$$

Из полученной матрицы можно сделать следующие выводы:

1) В течение рабочей недели пользователь x_1 взаимодействовал с сервером y_1 один раз, с сервером y_2 - четыре раза, а с сервером y_3 - не взаимодействовал.

2) Пользователь x_2 в этот период времени три раза взаимодействовал с сервером y_1 и по одному разу с серверами y_2 и y_3 ;

3) Из формулы (65) следует, что:

$$\arg \max_{\substack{1 \leq i \leq 2, \\ 1 \leq j \leq 3}} \{w_{ij}\} = (1, 2).$$

Это означает, что в течение 5 рабочих дней в системе «пользователи-серверы» максимальным образом было отмечено взаимодействие первого пользователя со вторым сервером.

Оценки математического ожидания и дисперсии числа пользователей и серверов, участвующих во взаимодействии элементов этой системы, вычисленные на основе данных Табл. 2 по формулам (66), (67), имеют вид:

$$\begin{aligned} \bar{N}_X^C &= \frac{10}{5} = 2, \quad \bar{N}_Y^C = \frac{8}{5} = 1,6. \\ D[N_X^C] &= 0, \quad D[N_Y^C] = \frac{2(1 - 1,6)^2 + 3(2 - 1,6)^2}{4} = 0,3. \end{aligned}$$

Из этих результатов следует, что в связи с тем, что $\bar{N}_X^C = 2$ и $D[N_X^C] = 0$, множество X^C является неслучайным и совпадает с множеством X , а в течение 5 рабочих дней во взаимодействии в среднем участвует 1,6 серверов из трех имеющихся.

Отметим, что в п. 16 приведенного выше алгоритма необходимо реализовать вывод результатов в форме Табл. 2 и в виде графов, представленных на Рис. 44-46.

В заключение этой Лекции укажем, что введенные выше неклассические виды отношений, которые нуждаются в дальнейших более детальных исследованиях, могут быть использованы при реализации имитационных моделей процессов и систем (см. Рис. 13).

Раздел III. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ИНФОРМАЦИОННЫХ СИСТЕМ

В существующей теории сложных систем для решения задач анализа рассматриваются следующие уровни их описания:

- функциональный уровень, отражающий процессы, протекающие в системе;
- морфологический уровень, который описывает состав элементов системы и их взаимосвязи (структуру системы);
- информационный уровень, отражающий состав циркулирующей в системе информации и ее использование.

Соответственно этому на практике применяют функциональные, структурные и информационные модели систем и их компонентов.

В данном разделе основное внимание уделяется первому виду модели, классификация которых представлена на Рис. 13.

12. Бинарные модели информационных систем

Рассмотрим предварительно ряд дополнительных определений из теории систем.

Под системой будем понимать множество взаимосвязанных элементов, действующих совместно с целью выполнения поставленной задачи.

Под анализом системы понимается процесс определения наиболее реальных способов выполнения целевых задач при удовлетворении заданных требований.

При исследовании системы одним из существенных мероприятий является формирование ее модели.

Моделью системы будем называть качественное и количественное представление системы, отражающее взаимосвязь и влияние факторов для решения задач ее анализа и синтеза.

Каждая модель должна описывать только существенные входные и выходные воздействия, а также учитывать влияние окружающих условий.

Классификация моделей систем описана в Лекции №4 и приведена на Рис. 13.

Рассмотрим основные моменты описания систем.

Состоянием системы называется множество существенных свойств, которыми она обладает в данный момент времени.

Характеристики свойств системы, которые поддаются непосредственному наблюдению, образуют множества входов X и выходов Y системы.

Входы, допускающие выбор конкретных воздействий, составляют множество $U \subset X$ управлений.

Функционированием системы называется изменение хотя бы одного из рассматриваемых свойств во времени.

Впервые теоретико-множественные модели систем были предложены в работе [Месарович М., Мако Д., Тахакара И.].

Пусть X – множество входных воздействий, Y – множество выходных воздействий (реакций) системы. Тогда моделью абстрактной системы называется объект вида:

$$S \subseteq X \times Y, \quad (69)$$

где S – бинарное отношение множеств X и Y .

Эта модель описывает систему, схематически представленную на Рис. 47.

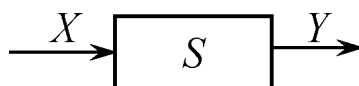


Рис. 47

Если ввести множество Z , которое описывает множество состояний системы, то абстрактная система может быть описана отношениями вида:

$$R_1 \subseteq X \times Z, \quad R_2 \subseteq Z \times Y.$$

В этом случае абстрактная модель системы (69) получается с помощью операции композиции и имеет следующий вид:

$$S = R_1 \circ R_2 \subseteq X \times Y.$$

Этими авторами было введено понятие функциональной системы, которая описывается с использованием понятия отображения множеств X и Y .

Модель функциональной системы в общем случае записывается как:

$$S: X \rightarrow Y, \quad (70)$$

где S – отображение указанных множеств.

Понятие отображения множеств является обобщением понятия функции и функциональной зависимости.

Функциональные модели систем обобщают понятие аналитических математических моделей, (см. Рис. 13) т.е. моделей, заданных в виде систем алгебраических, логических, дифференциальных, интегральных и конечно-разностных уравнений, описывающих процессы их функционирования.

При моделировании информационных систем (ИС) и процессов используются оба описанных выше подхода.

При этом множество X представляет собой совокупность документов, запросов, сообщений, сигналов и т.п., поступающих на вход ИС для последующей обработки. Множество Y составляют результаты функционирования системы в виде документов, видеоформ, сигналов и пр. Связь между этими множествами описывается бинарными отношениями, рассмотренными в Разделе II.

При использовании классических отношений выражение (69) представляет собой расчетную дискретную аналитическую статическую детерминированную модель рассматри-

ваемой ИС. При этом анализ системы в части ее взаимодействия с другими ИС можно формально осуществить с использованием теоретико-множественных и специальных операций над описывающими их отношениями.

Будем называть теоретико-множественные модели ИС, использующие такие отношения их бинарными моделями.

Простейшая бинарная модель ИС, построенная на основе выражения (69), имеет вид:

$$\mathfrak{M}_{\text{ис}} = \{X, Y, S\}.$$

Отметим, что любая бинарная модель ИС описывает систему как «черный ящик».

Если в качестве отношения S использовать случайное \mathcal{R} или стохастическое R^C отношения, то модель $\mathfrak{M}_{\text{ис}}$ превращается из детерминированной аналитической модели в вероятностную имитационную модель ИС. Заметим, что с помощью такой модели можно описать взаимодействие пользователей с некоторым сегментом Интернета и получить определенные числовые характеристики этого процесса.

При использовании динамического отношения S^t получаем дискретную имитационную динамическую численно разрешимую модель функционирования ИС на рассматриваемом интервале $[t_0, t_k]$ времени ее работы.

При формировании имитационной модели такой системы предлагается использовать следующий подход к построению матричной функции $A(t) = [a_{ij}(t)]_{n \times k}$.

Будем считать, что для описания каждого элемента $a_{ij}(t)$, $i = (\overline{1, n})$, $j = (\overline{1, k})$ задана совокупность значений t_r и τ_r , описывающая моменты времени перехода этой функции от значе-

ния «0» к значению «1» и время пребывания в состоянии «1», $r = (\overline{1, m_{ij}})$. На Рис. 48 приведена иллюстрация этого процесса.

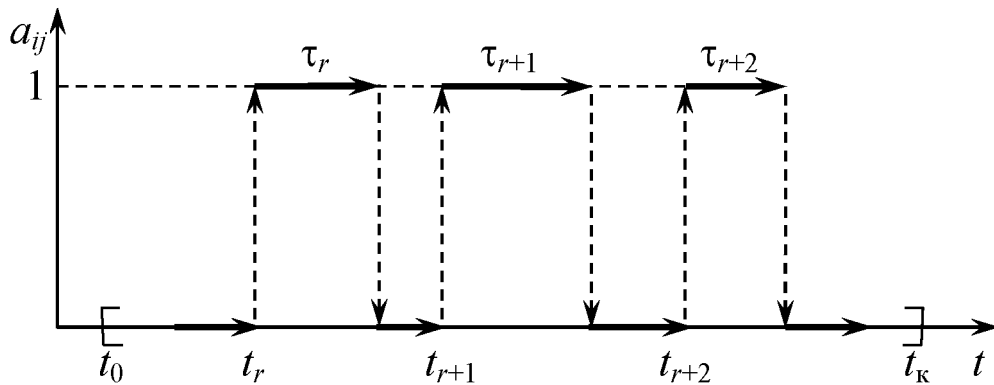


Рис. 48

Для формирования требуемых функций $a_{ij}(t)$, $t \in [t_0, t_k]$ предлагается использовать следующий алгоритм:

1. Ввод значений t_0 и t_k и шага Δt .
2. Ввод значений i, j, m_{ij} .
3. Ввод значений t_r и τ_r , $r = (\overline{1, m_{ij}})$.

4. Полагаем $t = t_0$, $r = 1$.

5. Проверка условия:

$$t_r < t < t_r + \tau_r.$$

6. Если это условие выполняется, то $a_{ij}(t) := 1$. В противном случае $a_{ij}(t) := 0$.

7. Полагаем $t := t + \Delta t$, проверяем выполнение условия:

$$t \leq t_k$$

и переходим к п. 5. В противном случае – переход к п. 10.

8. Полагаем $r := r + 1$ и проверяем условие:

$$r \leq m_{ij}.$$

Если оно выполняется – переход к п. 5. В противном случае – к п. 2.

9. Конец.

Как известно, работу информационной системы определяет реализованное в ее составе множество программ.

Пусть в рассматриваемой информационной системе имеется множество P программ обработки информации. Известно, что эти программы используют множество X исходных данных. В результате выполнения программ получается множество Y результатов обработки.

Если выбрать некоторое $x \in X$, то этот элемент данных может использоваться в определенной совокупности программ. С другой стороны программа $p \in P$ использует определённую совокупность данных из состава множества X . Таким образом, на декартовом произведении множеств $X \times P$ можно сформировать отношение:

$$R_1 \subseteq X \times P, \quad (71)$$

которое описывает связь типа «данные-программы».

Любая программа $p \in P$ формирует определённые подмножества результатов $y \in Y$. Кроме того, любой результат y в общем случае может быть получен с помощью нескольких программ, отличающихся друг от друга исходными данными, результирующей точностью и т.п. Это позволяет построить отношение вида:

$$R_2 \subseteq P \times Y \quad (72)$$

В этом случае более детальная модель некоторой ИС может быть записана как:

$$\mathfrak{M}_{\text{ис}} = \{P, X, Y, R_1, R_2\}. \quad (73)$$

Для исследования (анализа) ИС сформируем на основе этой модели некоторые вторичные отношения. Построим связь типа «данные-программы-результаты», используя операцию объединения отношений (71) и 72):

$$Q_1 = R_1 \cup R_2 \subseteq X \times P \times Y.$$

Данное отношение может быть представлено визуально в виде трёхслойного графа, который показывает связь между используемыми данными, программами и полученными результатами.

Определим множество программ для получения требуемых результатов. Для этого построим инверсию отношения (72) вида:

$$Q_2 = R_2^{-1} \subseteq Y \times P.$$

Сформируем отношение, описывающее состав исходных данных для получения конкретных результатов. Это отношение получается с использованием операции композиции отношений как:

$$Q_3 = R_1 \circ R_2 \subseteq X \times Y.$$

Построение отношений Q_1 , Q_2 и Q_3 при значительных мощностях множеств P , X и Y предлагается осуществлять с использованием компьютерных реализаций приведенных в Разд. II матричных операций над отношениями (71) и (72).

Если считать эти отношения динамическими отношениями R_1^t и R_2^t , $t \in [t_0, t_k]$, то параметрическое множество программ P^t можно рассматривать как множество состояний рассматриваемой ИС. При таком подходе получаем динамическую модель ИС вида:

$$\mathfrak{M}_{\text{ис}} = \{P^t, X^t, Y^t, R_1^t, R_2^t \mid t \in [t_0, t_k]\}.$$

Отметим, что математические модели, учитывающие современные концепции построения ИС, практически отсутствуют в существующей литературе. К таким основным концепциям относятся:

1. Широкое использование автоматизированных рабочих мест (АРМ), различных категорий пользователей ИС.

2. Организация данных в сложных ИС в форме распределённого банка данных (РБНД).

3. Наличие системы связи, позволяющей организовать обмен данными между взаимодействующими друг с другом АРМ, между АРМ и удалёнными базами данных и т.п.

4. Открытость ИС, которая подразумевает применение в системе прикладных программ, не зависящих от используемых платформ.

Сформируем с учетом этого модель ИС, предполагая, что она реализована в среде некоторого предприятия (организации).

Введем в рассмотрение следующие множества элементов ИС:

Λ – множество подразделений предприятия (организации);

Π – множество пользователей ИС.

A – множество АРМ и других технических средств, используемых в системе;

P – множество системных и прикладных программ системы;

B – множество локальных банков данных, содержащих всю необходимую информацию для функционирования системы.

На этих множествах зададим следующие первичные бинарные отношения:

- закрепление пользователей за конкретными АРМ:

$$R_1 \subseteq \Pi \times A;$$

- распределение программного обеспечения по АРМ:

$$R_2 \subseteq A \times P;$$

- связь программ и данных в ИС:

$$R_3 \subseteq P \times B;$$

- размещение АРМ по подразделениям предприятия:

$$R_4 \subseteq \Lambda \times A.$$

Тогда модель ИС можно представить в виде:

$$\mathfrak{M}_{\text{ис}} = \{ \Lambda, \Pi, A, P, B, R_1, R_2, R_3, R_4 \}. \quad (74)$$

Данная модель позволяет с использованием операций над этими отношениями получать дополнительную информацию об ИС. Построим с использованием модели (74) ряд вторичных отношений для проведения анализа рассматриваемой системы.

Описание взаимодействия АРМ с локальными банками данных может быть получено в виде следующего отношения:

$$Q_1 = R_2 \circ R_3 = (A \times P) \circ (P \times B) \subseteq A \times B.$$

Связь пользователей с банками данных можно представить в виде:

$$Q_2 = R_1 \circ R_2 \circ R_3 = (\Pi \times A) \circ (A \times P) \circ (P \times B) \subseteq \Pi \times B.$$

Взаимодействие между АРМ в процессе работы ИС, определяющее структуру электронного обмена данными между АРМ получаем в виде соотношения:

$$Q_3 = R_2 \circ R_3 \circ Q_1^{-1} = (A \times P) \circ (P \times B) \circ (A \times B)^{-1} \subseteq A \times A.$$

Взаимодействие банков данных, определяющее структуру РБНД, представляется в виде следующего отношения:

$$Q_4 = R_3^{-1} \circ R_2^{-1} \circ Q_1 = (B \times P) \circ (P \times A) \circ (A \times B) \subseteq B \times B.$$

Известно, что программы любой ИС, взаимодействуя между собой, определяют структуру системы обработки информации. Эта структура формально может быть описана отношением вида:

$$Q_5 = R_3 \circ Q_1^{-1} \circ R_2 \subseteq P \times P.$$

Пользователи ИС путём электронного обмена данными между АРМ и работы с общими данными взаимодействуют между собой в составе системы. Это взаимодействие может быть формально представлено как:

$$Q_6 = R_1 \circ Q_1 \circ Q_2^{-1} = (\Pi \times A) \circ (A \times B) \circ (B \times \Pi) \subseteq \Pi \times \Pi.$$

Структура электронного взаимодействия подразделений предприятия описывается отношением вида:

$$Q_7 = R_4 \circ Q_3 \circ R_4^{-1} \subseteq \Lambda \times \Lambda.$$

Матричное представление первичных отношений R_1, R_2, R_3, R_4 позволяет достаточно просто получать производные отношения $Q_1, Q_2, Q_3, \dots, Q_7$, которые наряду с задачами анализа ИС могут быть использованы для управления процессом её развития и формирования информационных технологий в системе.

Приведенные выше бинарные модели информационных систем представляют собой совокупности множеств и первичных отношений.

В общем случае задача оптимизации состава информационных систем формулируется как задача оптимального выбора подмножеств соответствующих элементов системы из априорно известных множеств этих компонентов.

Пусть синтезируемое множество X входных элементов системы (69) должно состоять из N типов элементов. Для каждого i -го элемента имеется множество $\Theta_i, i = (\overline{1, N})$ допустимых вариантов.

Тогда задача формирования парето-оптимальных вариантов множества X синтезируемой информационной системы представляется моделью вида:

$$\begin{aligned} C = \sum_{i=1}^N \sum_{j \in \Theta_i} c_{ij} \eta_{ij} \rightarrow \min; \quad E = \sum_{i=1}^N \sum_{j \in \Theta_i} e_{ij} \eta_{ij} \rightarrow \max; \\ \sum_{i=1}^N \sum_{j \in \Theta_i} a_{kij} \eta_{ij} \leq b_k, \\ k = (\overline{1, M}), \quad \eta_{ij} \in \{0, 1\}, \quad j \in \Theta_i, \quad i = (\overline{1, N}). \end{aligned} \quad (75)$$

Здесь булевские переменные x_{ij} определяют выбор j -го варианта реализации i -го входного элемента системы; c_{ij}, e_{ij} —

соответственно стоимость и эффективность использования j -го варианта i -го элемента в данной системе; a_{kij} — значение характеристики k -го вида j -го варианта i -го элемента; b_k — ограничение по k -й характеристике системы (характеристики объема, достоверности и т.п.).

Отметим, что использование таких моделей при синтезе ИС подразумевает выбор разработчиком ИС подходящего варианта на основе общепринятого в системотехнике компромисса между «стоимостью» и «эффективностью» системы.

Задачу, описываемую выражениями (75), можно решать при формировании множества Y выходных воздействий (реакций) ИС.

Сформулируем задачу синтеза первичного отношения S . Эта задача сводится к построению оптимальной булевой матрицы, описывающей формируемое (синтезируемое) отношение. Для обеспечения единого подхода к синтезу ИС здесь также предполагается использовать критерии стоимости и эффективности.

Пусть даны множества $X = \{x_i | i = (\overline{1, n})\}$, $Y = \{y_j | j = (\overline{1, m})\}$, состав которых определен из задач вида (75).

Требуется сформировать бинарное отношение (69). Как было показано выше, это отношение формально описывается матрицей $S = [s_{ij}]_{n \times m}$ с элементами $s_{ij} = 1$, если $(x_i, y_j) \in S$ и $s_{ij} = 0$, если $(x_i, y_j) \notin S$.

Обозначим через c_{ij} и e_{ij} соответственно стоимость и эффективность установления связи конкретных элементов $x_i \in X$ и $y_j \in Y$.

Считая величины s_{ij} переменными, удовлетворяющими условиям вида:

$$s_{ij} \in \{0; 1\}, \quad i = (\overline{1, n}), \quad j = (\overline{1, m}), \quad (76)$$

критерии оптимального синтеза отношения S представим как:

$$C = \sum_{i=1}^n \sum_{j=1}^m c_{ij} s_{ij} \rightarrow \min; \quad E = \sum_{i=1}^n \sum_{j=1}^m e_{ij} s_{ij} \rightarrow \max. \quad (77)$$

Условия того, что в формируемом отношении S будут отсутствовать изолированные элементы в множествах X и Y , можно записать в виде следующих неравенств:

$$\begin{aligned} \sum_{i=1}^n s_{ij} &\geq 1, \quad j = (\overline{1, m}); \\ \sum_{j=1}^m s_{ij} &\geq 1, \quad i = (\overline{1, n}). \end{aligned} \quad (78)$$

Пусть δ_{kij} – значение k -й характеристики, учитываемой при формировании связи между элементами $x_i \in X$ и $y_j \in Y$; Δ_k – ограничение по k -й характеристике системы. Например, если δ_{kij} – вероятность безотказной работы i -й программы на j -й ЭВМ и Δ_1 – вероятность безотказной работы системы, то ограничение на требуемую надежность процессов обработки данных можно представить как

$$\prod_{i=1}^n \prod_{j=1}^m \delta_{1ij}^{r_{ij}} \geq \Delta_1 \quad (79)$$

Отметим, что δ_{1ij} вычисляется как $\delta_{1ij} = p_i q_j$, где p_i – вероятность безотказной работы i -й программы; q_j – вероятность безотказной работы j -й ЭВМ системы.

В общем случае ограничения, описывающие K технических требований, учитываемых при синтезе рассматриваемой системы можно представить как:

$$\sum_{i=1}^n \sum_{j=1}^m \delta_{kij} s_{ij} \leq \Delta_k, \quad k = (\overline{1, K}) \quad (80)$$

Отметим, что ограничения вида (79) можно привести к линейной форме путём логарифмирования их правой и левой части.

Таким образом, для формального синтеза ИС предлагается решать известными методами двухкритериальные задачи булевского программирования (75) и (77), (76), (78), (80).

Отметим, что приведенные задачи можно использовать при формальном синтезе ИС, описываемых более сложными бинарными моделями.

13. Модели работы пользователей с банком данных информационной системы

Практика эксплуатации разнообразных информационных систем (ИС) показала, что их функционирование осуществляется при действии всевозможных случайных факторов. К таким факторам можно отнести случайные воздействия внешней среды, отказы и сбои аппаратных и программных средств системы, случайные потоки обращения к ним пользователей системы, случайные затраты времени на обработку и выдачу результатов их запросов и т.д.

Для описания и анализа систем различной природы, находящихся под действием случайных факторов в 60-70-х годах прошлого века была разработана теория массового обслуживания (ТМО). В силу ряда причин, в частности неразработанностью к тому времени основ кибернетики и весьма слабым применением ЭВМ в различных сферах деятельности общества, до сего времени ТМО по инерции практически не используется для решения задач моделирования анализа и синтеза ИС.

Рассмотрим краткую характеристику этой теории.

Объектом исследований и разработок задач ТМО является функционирование специального вида систем, называемых системами массового обслуживания (СМО). Примерами таких систем являются автоматические телефонные станции (АТС), ремонтные мастерские, билетные кассы, справочные бюро, магазины, парикмахерские, вычислительные центры коллективного пользования и т.п. [Вентцель Е.С.; Раскин Л.Г.; Кофман А., Крюон Р.].

Каждая СМО состоит из определённого числа обслуживающих единиц, которые называются каналами обслуживания. В качестве таковых могут фигурировать: линии связи АТС, персонал, разнообразные приборы и оборудование и т.п.

Системы массового обслуживания могут быть одноканальными или многоканальными.

На практике каждая СМО предназначена для обслуживания (выполнения) какого-то потока заявок («требований»), поступающих в неё в некоторые, в общем случае, случайные моменты времени. Обслуживание поступившей заявки продолжается некоторое (в общем случае, случайное) время, после чего канал освобождается и является готовым к принятию следующей заявки. Случайный характер потока заявок и времени их обслуживания приводит к тому, что в некоторые промежутки времени на входе СМО накапливается значительное число заявок, которые либо образуют очередь, либо покидают СМО не обслуженными. В другие периоды времени система может работать с недогрузкой или вообще простаивать.

Каждая система массового обслуживания, в зависимости от числа каналов и их производительности, а также от характера потока заявок, должна иметь определённую пропускную

способность, позволяющую ей успешно справляться с потоком заявок.

Предметом теории массового обслуживания, которая базируется на теории марковских случайных процессов [Раскин А.Г. и др.] является установление формальных зависимостей между характером потока заявок, числом каналов, их производительностью, правилами работы СМО и эффективностью обслуживания [Вентцель Е.С.].

В качестве характеристик эффективности обслуживания, в зависимости от условий задачи и целей исследования, могут применяться следующие показатели:

- среднее количество заявок, которое может обслужить СМО в единицу времени;
- средний процент заявок, получающих отказ и покидающих СМО не обслуженными;
- вероятность того, что поступившая заявка немедленно будет принята к обслуживанию;
- среднее время ожидания заявок в очереди на обслуживание;
- закон распределения времени ожидания обслуживания;
- среднее количество заявок, находящихся в очереди;
- закон распределения числа заявок в очереди;

Случайный характер потока заявок, а в общем случае и длительности обслуживания приводит к тому, что работа системы массового обслуживания будет описываться определённым случайным процессом, который изучается в теории массового обслуживания.

Системы массового обслуживания, в которых поток заявок поступает от конечного числа источников и возвращается к ним, называются замкнутыми СМО.

В сущности, любая СМО на практике имеет дело только с ограниченным числом источников заявок, но в ряде случаев число этих источников может быть весьма значительным, что позволяет пренебречь их имеющимся количеством. Например, поток вызовов на АТС от населения крупного города исходит, в сущности, от ограниченного числа абонентов, но это число так велико, что практически можно считать интенсивность потока заявок независимой от числа владельцев телефонов. Такие системы будем называть разомкнутыми СМО.

На Рис. 49 представлена классификация существующих систем массового обслуживания.



Рис. 49

В эту классификацию не включены СМО с определёнными дисциплинами обслуживания заявок в очереди («первый пришёл – первым обслужен», «нетерпеливые» заявки и т.п.).

Отметим, что наибольшее распространение на практике получили одно- и многоканальные, разомкнутые и замкнутые СМО с пуассоновскими законами распределения потоков заявок и показательными законами распределения времени их обслуживания [Вентцель Е.С.].

Рассмотрим простейшие вероятностные модели функционирования ИС, использующие математический аппарат описания замкнутых СМО.

Будем считать, что банк данных (БнД) информационной системы осуществляет работу с n пользователями, которые обращаются к нему с запросами в случайные моменты времени.

Практика использования БнД при функционировании разнообразных ИС показывает, что такие ситуации возникают случайным образом и образуют случайный поток «заявок» на их «обслуживание» (отработку) аппаратными и программными средствами банка. В связи с невозможностью точного прогнозирования содержания каждого из таких поступающих запросов к БнД в общем случае потребуются случайные затраты времени на их анализ и обслуживание.

Будем предполагать, что при поступлении некоторой заявки на обслуживание в момент времени, когда БнД занят обслуживанием заявки от другого пользователя, она ожидает обслуживания в очереди заявок на сервере ИС.

В связи с тем, что БнД (обслуживающее устройство) взаимодействует с n автоматизированными рабочими местами (АРМ) пользователей (источниками заявок), рассматриваемая система может быть отнесена к классу замкнутых систем массового обслуживания (СМО). Структура такой СМО приведена на Рис. 50.

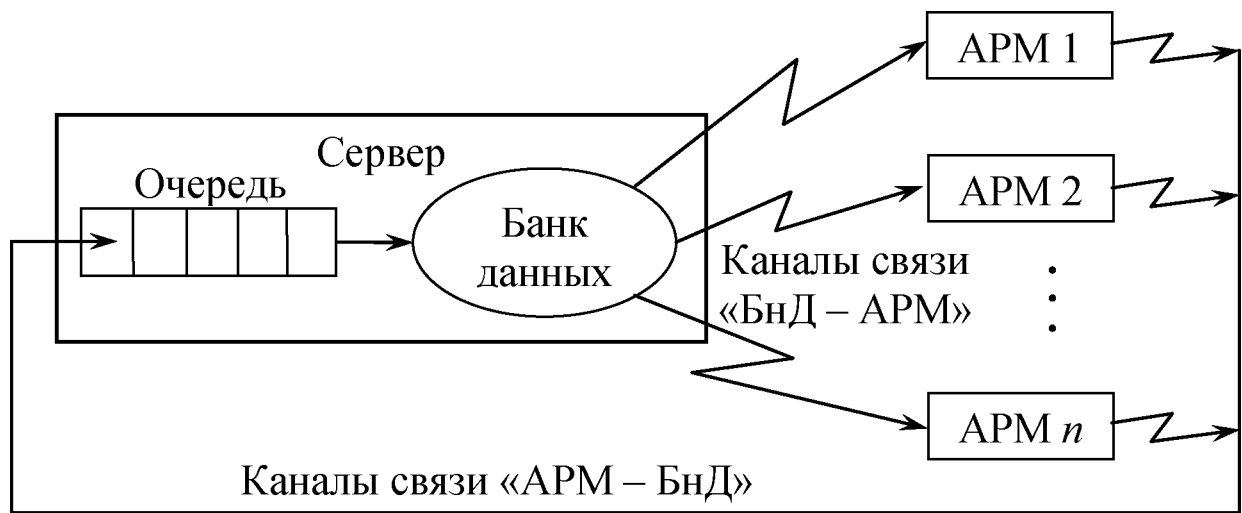


Рис. 50

Введем следующее множества состояний системы, в которых она может находиться в любой момент времени [Вентцель Е.С.]:

E_0 – БНД свободен, т.к. заявки от пользователей отсутствуют;

E_1 – БНД занят обслуживанием заявки, поступившей от одного пользователя;

E_2 – заявки поступили от двух пользователей, одна заявка обслуживается БНД, другая ожидает в очереди;

.....

E_n – заявки поступили от всех пользователей, одна – обслуживается, $(n - 1)$ заявка ожидает в очереди.

Граф связи состояний $E_0, E_1, E_2, \dots, E_n$ рассматриваемого марковского процесса с непрерывным временем и дискретным множеством состояний приведен на Рис. 51.

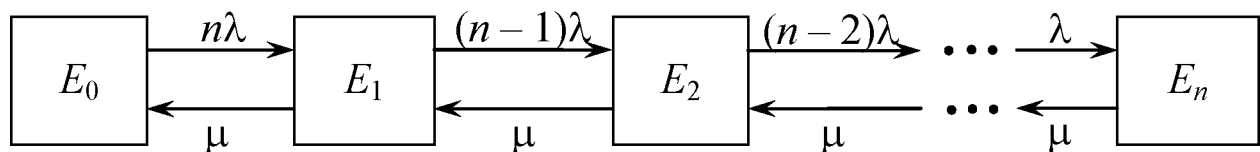


Рис. 51

На этом рисунке параметр λ означает интенсивность потока запросов от одного пользователя, μ — интенсивность его обслуживания банком данных системы.

Введем в рассмотрение вероятности $p_i(t)$ того, что в момент времени t система находится в состоянии E_i , $i = (\overline{0, n})$.

Система дифференциальных уравнений, описывающих динамику изменения этих вероятностей, построенная по графу, представленному на Рис. 51, имеет следующий вид [Вентцель Е.С.]:

$$\begin{aligned} \dot{p}_0 &= -n\lambda p_0(t) + \mu p_1(t); \\ \dot{p}_1 &= -[(n-1)\lambda + \mu]p_1(t) + n\lambda p_0(t) + \mu p_2(t); \\ \dot{p}_2 &= -[(n-2)\lambda + \mu]p_2(t) + (n-1)\lambda p_1(t) + \mu p_3(t); \\ &\dots\dots\dots \\ \dot{p}_i &= -[(n-i)\lambda + \mu]p_i(t) + (n-i+1)\lambda p_{i-1}(t) + \mu p_{i+1}(t); \\ &\dots\dots\dots \\ \dot{p}_n &= -\mu p_n(t) + \lambda p_{n-1}(t). \end{aligned} \tag{81}$$

Начальные условия для этой системы записываются как:

$$p_0(0) = 1, \quad p_i(0) = 0, \quad i = (\overline{1, n}). \tag{82}$$

Вероятности $p_i(t)$, определяемые из решения задачи Коши (81), (82), должны удовлетворять условию:

$$\sum_{i=0}^n p_i(t) = 1, \quad t \geq 0, \tag{83}$$

которое означает, что система в любой момент времени должна находиться в одном из состояний E_i , $i = (\overline{0, n})$.

Соотношения (81)-(83) описывают расчетную непрерывную аналитическую линейную динамическую вероятностную численно разрешимую модель (см. Рис. 13) взаимодействия пользователей с БНД ИС. Эти выражения являются отображением S в обозначении функциональной системы (70).

Рассматриваемая СМО может функционировать в следующих режимах:

1) *неустановившийся режим*, описываемый моделью (81)-(83);

2) *стационарный режим*, который характеризуется тем, что все производные $\dot{p}_i(t) = 0$ и $\lim_{t \rightarrow \infty} p_i(t) = p_i^* = \text{const}$, $i = (\overline{0, n})$.

Используя эти условия, получаем из системы уравнений (81) выражения вида:

$$\begin{aligned} & -n\lambda p_0^* + \mu p_1^* = 0; \\ & -[(n-1)\lambda + \mu]p_1^* + n\lambda p_0^* + \mu p_2^* = 0; \\ & \dots\dots\dots \\ & -[(n-i)\lambda + \mu]p_i^* + (n-i+1)\lambda p_{i-1}^* + \mu p_{i+1}^* = 0; \\ & \dots\dots\dots \\ & -\mu p_n^* + \lambda p_{n-1}^* = 0. \end{aligned}$$

Полученные соотношения наряду с условием (83), которое для установившегося режима записывается как:

$$\sum_{i=0}^n p_i^* = 1,$$

представляют собой статическую вероятностную модель работы системы «Пользователь – БНД».

Решения приведенной выше системы линейных алгебраических уравнений $(n+1)$ -го порядка имеют вид [Вентцель Е.С.]:

$$p_1^* = n\rho p_0^*; p_2^* = n(n-1)\rho^2 p_0^*; \dots; p_n^* = n!\rho^n p_0^*, \quad (84)$$

где $\rho = (\lambda/\mu)$ – коэффициент загрузки системы. Входящая в эти выражения вероятность p_0^* определяется по формуле вида:

$$p_0^* = [1 + n\rho + n(n-1)\rho^2 + \dots + n!\rho^n]^{-1}. \quad (85)$$

Отметим, что выражения (84), (85), которые описывают статическую аналитически разрешимую модель, могут быть использованы в случае, когда ИС функционирует достаточно большое время, за которое система «БНД – Пользователи» перешла в установившийся стационарный режим функционирования. Типичное поведение вероятностей $p_i(t)$, $t \geq 0$, $i = (\overline{0, n})$ представлено на Рис. 52.

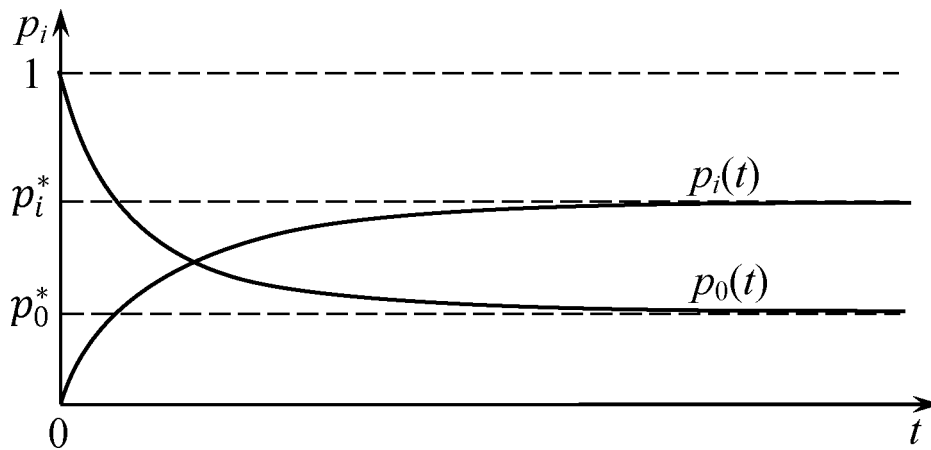


Рис. 52

Отметим, что условие наличия в СМО стационарного (установившегося) режима работы имеет вид [Вентцель Е.С.]:

$$\lambda < \mu \text{ или } \rho < 1.$$

При практическом использовании предлагаемых моделей в начале эксплуатации ИС необходимо численно решать задачу (81)-(83). При достаточно длительном времени эксплуатации ИС используются выражения (84) и (85).

В последнем случае при решении задач анализа ИС вычисляются следующие характеристики эффективности процесса обслуживания пользователей [Вентцель Е.С.]:

1) вероятность того, что в текущий момент времени БНД ИС занят обслуживанием поступившего запроса, вычисляемая по формуле:

$$P_{\text{зан}} = 1 - p_0^*,$$

2) абсолютная пропускная способность БНД системы, имеющая смысл интенсивности выходного потока обслуженных запросов:

$$A = (1 - p_0^*)\mu,$$

3) среднее число пользователей ИС, требующих обслуживания в текущий момент времени:

$$\bar{W} = n - \frac{1 - p_0^*}{\rho},$$

4) среднее количество пользователей, ожидающих обработки своих запросов (средняя длина очереди):

$$\bar{r} = n - \frac{(1 - p_0^*)(1 + \rho)}{\rho}$$

5) коэффициент ожидания пользователей их обслуживания банком данных системы:

$$K = \frac{\bar{r}}{n}$$

Для анализа работы информационной системы, необходимо определить значения приведенных показателей эффективности при имеющихся значениях параметров n и интенсивностей λ и μ , которые получаются путем обработки статистики собираемой соответствующими программами-мониторами СУБД банка данных.

При достаточно большом числе пользователей БНД предлагается использовать формализм разомкнутой одноканальной СМО [Вентцель Е.С.].

В этом случае имеем следующие показатели эффективности обслуживания БНД запросов пользователей ИС:

1) среднее число запросов в очереди на их обслуживание:

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

2) среднее время ожидания обслуживания:

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

3) среднее время ожидания запроса в очереди и его обслуживание:

$$\hat{t} = \frac{1}{\mu(1-\rho)},$$

где $\rho = \lambda/\mu$ – коэффициент загрузки системы.

Входящие в приведенные выше модели оценки параметров λ и μ , исходя из свойств экспоненциальных потоков, могут быть статистически определены следующим образом.

Пусть τ_k – интервал времени между поступлением в систему $(k-1)$ -го и k -го запросов, $k = (\overline{1, K})$. Тогда среднее значение интервала времени между поступлением запросов от пользователей по имеющейся выборке определяется как:

$$\bar{\tau} = \frac{1}{K} \sum_{k=1}^K \tau_k.$$

Для пуассоновского потока заявок имеем, что [Вентцель Е.С.]:

$$\lambda = \frac{1}{\bar{\tau}}.$$

Пусть θ_r – время обслуживания r -го запроса, $r = (\overline{1, R})$, где R – объем имеющейся выборки результатов мониторинга работы БНД ИС. Тогда средние затраты времени на обработку одного запроса будут равны:

$$\bar{\theta} = \frac{1}{R} \sum_{r=1}^R \theta_r.$$

Интенсивность μ вычисляется по формуле вида [Вентцель Е.С.]:

$$\mu = 1/\bar{\theta}.$$

Пример.

Пусть в процессе наблюдений за работой некоторой информационно-справочной системы получены следующие средние значения интервалов времени между поступлением запросов и их выполнением:

$$\bar{\tau} = 4 \text{ мин}; \quad \bar{\theta} = 0,4573 \text{ мин.}$$

Тогда оценки интенсивности поступления запросов к БНД и их выполнения, а также коэффициента загрузки системы будут равны:

$$\lambda = 0,25 \text{ запр./мин}; \quad \mu = 2,186 \text{ запр./мин}; \quad \rho = \frac{\lambda}{\mu} = 0,1143.$$

Вычислим по формуле (85) при $n = 6$ значение:

$$\begin{aligned} p_0^* &= \frac{1}{1 + 6 \cdot 0,1143 + 6 \cdot 5 \cdot 0,1143^2 + \dots + 6! \cdot 0,1143^6} = \\ &= \frac{1}{2,3349} = 0,428. \end{aligned}$$

Определим характеристики эффективности обслуживания пользователей системы.

Вероятность того, что в текущий момент времени БНД занят обслуживанием заявки:

$$P_{\text{зан}} = 1 - 0,428 = 0,572.$$

Абсолютная пропускная способность БНД:

$$A = (1 - 0,428) \cdot 2,186 = 1,252.$$

Среднее число пользователей, требующих обслуживания в текущий момент времени:

$$\bar{W} = 6 - \frac{1 - 0,428}{0,1143} = 1,002.$$

Среднее количество пользователей, ожидающих обслуживания (средняя длина очереди):

$$\bar{r} = 6 - \frac{(1 - 0,428)(1 + 0,1143)}{0,1143} = 0,428.$$

Коэффициент ожидания запросов на их обслуживания БНД ИС:

$$K = \frac{0,428}{6} = 0,071.$$

Самостоятельная работа.

Провести анализ изменения показателей эффективности работы БНД в процессе его внедрения в течение 8 недель с поэтапным расширением числа пользователей, равным 6,12; 18,24; 30,36; 42,48. Результаты расчетов привести в Табл. 3.

Таблица 3.

неделя	n	λ	μ	ρ	p_0	$P_{\text{зан}}$	A	\bar{W}	\bar{r}	K
1	6	0,25	2,186	0,1143	0,428	0,572	1,252	0,1143	0,428	0,071
2	12									
3	18									
4	24									
5	30									
6	36									
7	42									
8	48									

Сделать выводы об изменении показателей функционирования БНД с ростом числа его пользователей.

Отметим, что приведенные выше модели могут быть использованы для описания процессов работы пользователей в архитектурах «клиент-сервер» и «облачные вычисления».

14. Модель корпоративной информационной системы

Корпоративные информационные системы (КИС) получили в настоящее время широкое распространение как форма информатизации деятельности крупных предприятий, объединений, организаций и т.п.

Достаточно подробные структуры, функции и классификация существующих КИС приведены в монографии [Зиновьев П.А., Мейко А.В., Моисеев В.С.]. Вместе с тем можно отметить отсутствие у нас в стране и за рубежом достаточно разработанной современной теории КИС.

На наш взгляд одной из важнейших задач такой теории должна быть задача моделирования деятельности КИС. На основе таких моделей в последующем можно формулировать и решать задачи анализа и совершенствования действующих систем и синтеза разрабатываемых КИС новых предприятий, объединений и т.п.

Заметим, что модели КИС в доступной литературе практически отсутствуют.

Будем считать, что корпорация (предприятие, организация, объединение и т.п.) представляется множеством $\Lambda = \{0, 1, 2, \dots, m\}$ узлов обработки информации, в качестве которых выступают их департаменты, подразделения, отделы, службы и т.п.

Взаимосвязи узлов КИС будем описывать отношением вида:

$$R \subseteq \Lambda \times \Lambda, \quad (86)$$

которое представляется некоторым двудольным ориентированным графом, имеющим $2(m + 1)$ вершин.

Дуги этого графа определяют направления передачи информации в КИС.

В связи с большим многообразием циркулирующей в системе информации различного содержания и назначения будем считать, что затраты на ее обработку в узлах КИС являются случайными величинами. Кроме этого, за счет непрогнозируемых ситуаций в современной рыночной экономике появление новых данных в КИС и необходимости их обработки происходит в случайные моменты времени ее функционирования.

Все это позволяет рассматривать узлы КИС в общем случае как многоканальные системы массового обслуживания (СМО), обрабатывающие транзакции, поступающие им по каналам связи. После обработки последние передаются в виде транзакций другим узлам КИС в соответствии с приведенным выше отношением (86). При этом в процессе функционирования КИС адресаты передачи данных являются случайными.

Системы с описанными выше свойствами в литературе по теории массового обслуживания [Кофман А., Крюон Р.; Кузин Л.Т. и др.] носят название сетей СМО.

По существующей классификации таких сетей выделяют разомкнутые и замкнутые сети СМО.

В разомкнутых сетях один из ее узлов (вершин графа сети) выступает в качестве источника заявок на обслуживание и потребителя обработанных СМО сети заявок. В сетях второго типа заявки формируются и «потребляются» в составляющих их СМО.

В связи с тем, что любая КИС взаимодействует с внешней средой, т.е. является открытой системой, при ее моделировании будем использовать формализм разомкнутых сетей СМО.

В этом случае нулевой узел множества Λ , в качестве которого обычно выступает управляющий орган корпорации, бу-

дем считать источником заявок (транзакций) и потребителем результатов их обработки.

Предположим, что транзакции, выходящие из i -го узла с вероятностью θ_{ij} поступают в j -й узел системы, $i, j = (\overline{0, m})$. При этом из источника в j -й узел транзакции передаются с вероятностью θ_{0j} , $j \in (\overline{0, m})$ и возвращаются с вероятностями θ_{i0} , $i \in (\overline{0, m})$. Описание возможного взаимодействия узлов КИС производится с помощью матрицы передач транзакций:

$$T = \begin{bmatrix} 0 & \theta_{01} & \theta_{02} & \dots & \theta_{0m} \\ \theta_{10} & \theta_{11} & \theta_{12} & \dots & \theta_{1m} \\ \dots & \dots & \dots & \dots & \dots \\ \theta_{m0} & \theta_{m1} & \theta_{m2} & \dots & \theta_{mm} \end{bmatrix}. \quad (87)$$

В этой матрице элемент $\theta_{00} = 0$, т.к. считается, что циркулирование транзакций в источнике отсутствует.

Элементы матрицы (87) должны удовлетворять следующим требованиям:

$$0 \leq \theta_{ij} \leq 1; \quad \sum_{j=0}^m \theta_{ij} = 1, \quad i, j = (\overline{0, m}). \quad (88)$$

Учет соотношений (87) и (88) превращает отношение (86) в стохастическое отношение R^c (см. Лекцию №11).

Будем считать, что сеть узлов КИС работает в установившемся (стационарном) режиме функционирования.

Предполагая, что среднее число транзакций в единицу времени, покидающих j -й узел, равно среднему числу поступивших в него транзакций в единицу времени от других узлов сети, запишем выражения вида:

$$\lambda_j = \sum_{i=0}^m \lambda_i \theta_{ij}, \quad j = (\overline{0, m}), \quad (89)$$

где λ_i – интенсивность выходного потока транзакций i -го узла, $i = (\overline{0, m})$.

Исходя из этого предположения, рассматриваемые сети получили название линейных стохастических сетей СМО [Кофман А., Крюон Р.].

Будем считать, что значение интенсивности λ_0 источника транзакций является заданной величиной. Тогда выражения (89) можно рассматривать как систему линейных алгебраических уравнений $(m + 1)$ -го порядка. Эта система может рассматриваться как непрерывная аналитическая теоретическая линейная статическая вероятностная модель КИС (см. Рис. 13).

Система (89), имеющая ранг, равный m , может быть решена аналитически или известными численными методами [Вержбицкий В.М.].

Пусть $\lambda_j^{(0)}$ – корни этой системы, $j = (\overline{1, m})$. Введем в рассмотрение характеристики сети, которые будут использованы при ее количественном анализе. Количество прохождений транзакций через j -й узел сети прежде, чем попасть в источник, определяется как:

$$\alpha_j = \frac{\lambda_j}{\lambda_0}, \quad j = (\overline{1, m}). \quad (90)$$

Наиболее и наименее нагруженные узлы сети определяются из соотношений вида:

$$\begin{aligned} j_{\max} &= \arg \max \{ \alpha_j \}, \quad j \in (\overline{1, m}); \\ j_{\min} &= \arg \min \{ \alpha_j \}, \quad j \in (\overline{1, m}). \end{aligned} \quad (91)$$

Обозначим через τ_j среднее время обработки (пребывания) транзакции в j -м узле сети, $j \in (\overline{1, m})$. Отсюда среднее

время обработки информации, поступающей в сеть рассматриваемой КИС, вычисляется по следующей формуле:

$$\tau = \sum_{j=1}^m \alpha_j \tau_j. \quad (92)$$

Доказано, что, если в каждом узле сети время обслуживания распределено по показательному закону с параметрами μ_j , $j = (\overline{1, m})$, то выходные потоки заявок каждого узла будут пуассоновскими потоками [Кузин Л.Т.; Кофман А., Крюон Р.]. Такие сети СМО называются линейными показательными сетями массово обслуживания. Для этого класса сетей будет существовать установившийся режим их работы при выполнении для каждого узла условия:

$$\rho_j = \frac{\lambda_j^{(0)}}{\mu_j} < 1, \quad j = (\overline{1, m}).$$

Известно, что для показательного закона имеет место следующее равенство [Вентцель Е.С.]:

$$\tau_j = \frac{1}{\mu_j}, \quad j = (\overline{1, m}).$$

Тогда формула (92) с учетом этого и выражения (90) примет вид:

$$\tau = \frac{1}{\lambda_0} \sum_{j=1}^m \left(\frac{\lambda_j^{(0)}}{\mu_j} \right) = \frac{1}{\lambda_0} \sum_{j=1}^m \rho_j. \quad (93)$$

Пример.

Рассмотрим простейшую КИС, состоящую из множества узлов $\Lambda = \{0; 1; 2\}$, граф связи [Кузин Л.Т.] которых приведен на Рис. 53.

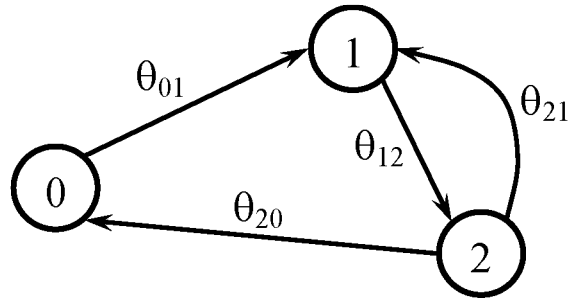


Рис. 53

Здесь узел «0» является источником транзакций. Матрица их передач (87) записывается как:

$$T = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{bmatrix} 0 & \theta_{01} & 0 \\ 0 & 0 & \theta_{12} \\ \theta_{20} & \theta_{21} & 0 \end{bmatrix} \end{matrix}$$

С учетом условий (88) имеем, что:

$$\theta_{01} = 1; \quad \theta_{12} = 1; \quad \theta_{20} = \theta; \quad \theta_{21} = 1 - \theta.$$

где θ – заданное значение вероятности завершения обработки транзакции в рассматриваемой сети.

С учетом вида матрицы T выражения (89) примут следующий вид:

$$\lambda_0 = \lambda_2 \theta_{20} = \lambda_2 \theta;$$

$$\lambda_1 = \lambda_0 \theta_{01} + \lambda_2 \theta_{21} = \lambda_0 + \lambda_2 (1 - \theta);$$

$$\lambda_2 = \lambda_1 \theta_{12} = \lambda_1.$$

Из первого и третьего уравнений этой системы получаем решения вида:

$$\lambda_2^{(0)} = \lambda_1^{(0)} = \frac{1}{\theta} \lambda_0.$$

Тогда среднее число проходов транзакций через узлы 1 и 2, согласно формуле (90), будет равно:

$$\alpha_1 = \alpha_2 = \frac{1}{\theta}.$$

Пусть μ_1 и μ_2 – интенсивности обработки транзакций в этих узлах. Тогда по формуле (93) среднее время обработки информации в сети определяется как:

$$\tau = \frac{1}{\lambda_0} \left(\frac{\lambda_0}{\mu_1 \theta} + \frac{\lambda_0}{\mu_2 \theta} \right) = \frac{\mu_1 + \mu_2}{\mu_1 \mu_2 \theta}.$$

Условия существования в сети установившегося режима обработки транзакций имеют вид:

$$\frac{\lambda_0}{\theta \mu_1} < 1; \quad \frac{\lambda_0}{\theta \mu_2} < 1.$$

Отсюда значение вероятности θ , обеспечивающее наличие такого режима, определяется оценкой вида:

$$\theta > \max \left\{ \frac{\lambda_0}{\mu_1}; \frac{\lambda_0}{\mu_2} \right\}.$$

Пусть имеем следующие условные числовые значения параметров поступления и обработки транзакций:

$$\lambda_0 = 0,1 \text{ мин}^{-1}; \quad \mu_1 = 4 \text{ мин}^{-1}; \quad \mu_2 = 8 \text{ мин}^{-1}.$$

Тогда значение вероятности θ завершения обработки должно удовлетворять неравенству:

$$\theta > \max \left\{ \frac{0,1}{4}; \frac{0,1}{8} \right\} = \max \{0,0250; 0,0125\} = 0,025.$$

Для выбранной вероятности $\theta = 0,026$ параметры α_1 , α_2 и τ примут следующие значения:

$$\alpha_1 = \alpha_2 = 38,46; \quad \tau = 14,423 \text{ мин.}$$

Отсюда можно сделать вывод, что среднее время обработки поступившей в систему транзакции равно $\sim 14,5$ мин. При этом транзакции проходят через одинаково нагруженные узлы 1 и 2 порядка 38 раз. Причиной последнего является достаточно большое значение вероятности $\theta_{21} = 1 - 0,026 = 0,974$ возврата транзакции из узла 2 на повторную обработку в узел 1.

Полагая значение $\theta = 1$, имеем:

$$\alpha_1 = \alpha_2 = 1; \quad \tau = 0,375 \text{ мин.}$$

Это означает, что каждая поступившая из источника транзакция проходит перед возвратом в него по одному разу через узлы 1 и 2. Среднее время ее пребывания в сети равно 0,375 мин.

Для сети с множеством узлов $\Lambda = \{0, 1, 2, 3\}$ и матрицей (87) [Кофман А., Крюон Р.]:

$$T = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0,5 & 0,5 \\ 0 & 0,25 & 0 & 0,75 \\ 0,5 & 0 & 0,25 & 0,25 \end{bmatrix} \end{matrix}$$

система уравнений (89) записывается как:

$$\lambda_0 = \lambda_3 \cdot 0,5;$$

$$\lambda_1 = \lambda_0 \cdot 1 + \lambda_2 \cdot 0,25;$$

$$\lambda_2 = \lambda_1 \cdot 0,5 + \lambda_3 \cdot 0,25;$$

$$\lambda_3 = \lambda_1 \cdot 0,5 + \lambda_2 \cdot 0,75 + \lambda_3 \cdot 0,25.$$

Аналитическое решение этой системы имеет вид:

$$\lambda_1^{(0)} = 1,286\lambda_0; \quad \lambda_2^{(0)} = 1,143\lambda_0; \quad \lambda_3^{(0)} = 2\lambda_0.$$

Отсюда, согласно формуле (90), средние значения числа прохождения транзакций через узлы сети будут равны:

$$\alpha_1 = 1,286; \quad \alpha_2 = 1,143; \quad \alpha_3 = 2.$$

Из выражений (91) следует, что наиболее нагруженным узлом сети является узел 3, а наименее нагруженным – узел 2.

Если предположить, что:

$$\lambda_0 = 2 \text{ мин}^{-1}; \quad \mu_1 = 4 \text{ мин}^{-1}; \quad \mu_2 = 3 \text{ мин}^{-1}; \quad \mu_3 = 6 \text{ мин}^{-1},$$

то коэффициенты загрузки узлов сети вычисляются как:

$$\rho_1 = \frac{1,286 \cdot 2}{4} = 0,643; \quad \rho_2 = \frac{1,143 \cdot 2}{3} = 0,762; \quad \rho_3 = \frac{2 \cdot 2}{6} = 0,666.$$

Тогда среднее время пребывания транзакции в сети, вычисляемое по формуле (93), будет равно:

$$\tau = \frac{1}{2}(0,643 + 0,762 + 0,666) = 1,036 \text{ мин.}$$

Оценки вероятностей $\theta_{ij}, i, j = (\overline{0, m})$, составляющих матрицу T передач транзакций, могут быть приближенно получены из предварительных наблюдений за работой каждого узла сети на определенном интервале времени функционирования КИС.

Пусть l_{ij} – среднее число транзакций, пересылаемых из узла $i = (\overline{0, m})$ в узел $j = (\overline{0, m})$. Тогда, согласно условиям (88), искомые оценки вероятностей передач транзакций вычисляются по формулам вида:

$$\theta_{ij} = \frac{l_{ij}}{\sum_{j=0}^m l_{ij}}, \quad i, j = (\overline{0, m}).$$

С практическим применением приведенной выше модели для описания работы корпоративной системы одного из крупнейших районных управлений почтовой связи г. Казани можно ознакомиться в статье [Гинатуллин И.А., Моисеев В.С.].

В заключение лекции отметим, что подробное и достаточно простое описание теории линейных стохастических сетей массового обслуживания содержится в монографии [Кофман А., Крюон Р.].

Раздел IV. МОДЕЛИ ПРИКЛАДНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Рассматриваемый класс технологий является основой любой информационной системы.

Анализ существующих работ по информатике показал практически полное отсутствие формализации общего описания прикладной информационной технологии и задач ее оптимального формирования.

В данном разделе предлагается один из подходов к решению этой важной для прикладной информатики проблемы.

15. Модели прикладной информационной технологии

В настоящее время на практике широко используется следующие виды общих информационных технологий (ИТ):

- Обработки текстов.
- Обработки изображения.
- Обработки звука.
- Мультимедийные технологии.
- WEB – технологии.

Эти технологии носят название общих, так как они практически не связаны с предметной областью решаемых с их помощью задач.

В 90-х годах прошлого века сформировалось понятие прикладных ИТ, ориентированных на решение конкретных научно-технических, инженерных, экономических и организационных задач. В составе каждой из таких технологий можно выделить следующие основные компоненты:

1. Вычислительная технология, основанная на численных методах и алгоритмах решения прикладной задачи.

2. Человеко-машинная (интерактивная) технология сбора, хранения, передачи и обработки данных.

3. Технология принятия решения.

Основу прикладных ИТ составляет технология 1-го вида (Рис. 54). Общие ИТ могут являться составными частями второго и третьего вида технологий.

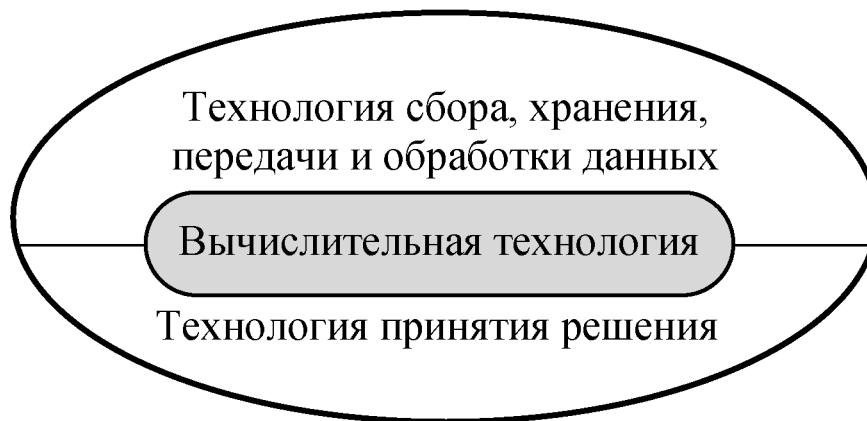


Рис. 54

В сложных научно-технических задачах при реализации вычислительной компоненты соответствующих прикладных ИТ используются современные численные методы проведения расчетов и вычислительных экспериментов. Например, при формировании прикладной ИТ моделирования работы пользователей с БНД в неустановившемся режиме его функционирования (см. Лекцию №13) для решения системы дифференциальных уравнений (81) используется один из численных методов ее интегрирования (Эйлера, Рунге-Кутта и др.) [Вержбицкий В.М.].

В других видах задач такие технологии реализуют соответствующие методики требуемых расчетов. При этом принятие решений по их результатам проводит пользователь соответствующей прикладной ИТ на основе своего опыта и целей ее применения.

Отдельный вид перспективных прикладных ИТ составляют технологии оптимизации проектных, управленческих и организационных решений.

Отметим, что в современных ИС крайне слабо осуществляется информатизация этой важнейшей проблемы. В лучшем случае в них имеются либо средства для проведения диалоговых расчетов по принципу «что будет, если ...», либо вывода результатов во всевозможных графических формах (OLAP-системы).

Современная технология принятия решения включает в себя формирование множества допустимых вариантов решения и выбор лицом, принимающим решение (ЛПР), конкретного варианта. При этом последнее может сопровождаться дополнительными расчётами, формированием графических образов вариантов и т.п.

Например, при формировании оптимальных по Парето решений задач многокритериальной оптимизации [Подиновский В.В., Ногин В.Д.] на завершающей стадии принятия решения могут быть использованы следующие процедуры:

1. Вывод таблицы вариантов в пространствах решений и критериев с дополнительной информацией в удобной для лица, принимающего решение, форме по каждой точке сформированного множества вариантов.

Для двухкритериальной задачи такое множество в пространстве критериев (W_1, W_2) приведено на Рис. 55.

Реализация этой процедуры предполагает, что в дополнении к значениям $(W_1^{(i)}, W_2^{(i)})$, $i = (\overline{1, n})$ выводится в числовом или графическом виде дополнительная информация, упрощающая лицу, принимающему решение, выбор из них компромиссного решения.

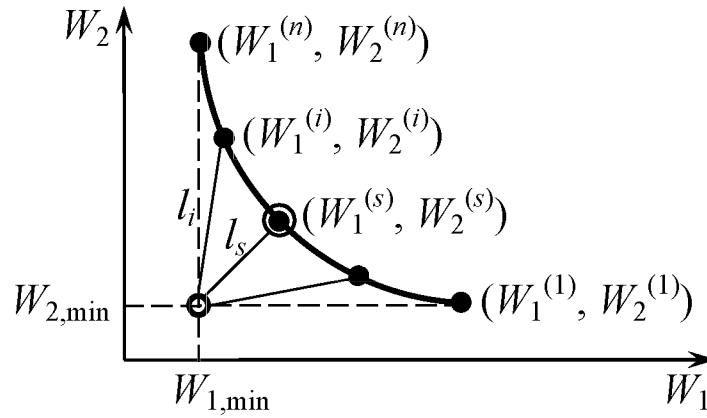


Рис. 55

2. Формальный выбор «наилучшей s -й точки» в множестве Парето, согласно следующему правилу:

$$s = \arg \min_{i \in (1, n)} \{l_i\} = \\ = \arg \min_{i \in (1, n)} \left\{ \sqrt{(W_1^{(i)} - W_{1, \min})^2 + (W_2^{(i)} - W_{2, \min})^2} \right\}.$$

Дадим краткую характеристику наиболее распространенного в настоящее время подхода к построению оптимальных по Парето (неулучшаемых) решений, основанного на применении линейной свертки целевых функций.

Пусть требуется найти вектор $x = (x_1, x_2, \dots, x_m)$, доставляющий минимальные значения целевым функциям:

$$W_1(x) \rightarrow \min_x; \quad W_2(x) \rightarrow \min_x.$$

Известно, что эта задача не имеет единственного решения.

Для построения паретооптимального множества векторов x предлагается решать задачу минимизации линейной свертки целевых функций вида:

$$L(x, \alpha) = \alpha W_1(x) + (1 - \alpha) W_2(x) \rightarrow \min_x,$$

где $\alpha \in [0; 1]$ – параметр свертки.

Задаваясь сеткой значений этого параметра:

$$0 = \alpha_1 < \alpha_2 < \alpha_3 < \dots < \alpha_n = 1$$

и решая при каждом α_i однокритериальную задачу минимизации $L(x, \alpha_i) \rightarrow \min_x$, получаем множество оптимальных по Парето векторов $x^{(i)}$ и значений критериев $W_1^{(i)}, W_2^{(i)}, i = (\overline{1, n})$ решаемой задачи.

Из этого множества ЛПР на основании его опыта, интуиции и дополнительных расчетов может выбирать требуемое в конкретной ситуации решение $x^{(s)}, s \in (\overline{1, n})$.

Отметим, что в многокритериальных задачах дискретной оптимизации такой подход не позволяет получать все точки множества паретооптимальных решений. Однако, как показала практика, получаемое с его помощью количество вариантов бывает вполне достаточно для их анализа и принятия соответствующего решения.

Любая прикладная ИТ включает в себя совокупность процедур выполняемых действий, а также применяемые при ее реализации программные, аппаратные средства и наборы данных.

В современных прикладных ИТ выделяются следующие группы процедур:

- Сбор требуемой информации и её регистрация в соответствующей базе данных (БД).
- Получение информации по запросу конкретного пользователя или прикладной программы.
- Решение конкретной прикладной задачи пользователя.
- Передача результатов решения задачи лицу, принимающему решения.
- Процедуры, связанные с процессом выбора конкретного решения.

Обобщённая схема связи процедур прикладной ИТ приведена на рис. 56.



Рис. 56

В предлагаемой схеме реализованы следующие особенности:

1. При доступе пользователя к информационной системе проверяются его полномочия на работу с определёнными данными, программами и оборудованием.

2. Операции по сбору, контролю, вводу и передаче информации в удалённые базы данных распределённого банка данных (РБНД) могут выполняться автономно по мере появления новой информации.

3. После формирования запроса пользователя к системе проводится идентификация запроса на предмет допуска пользователя к определённой информации.

4. Требуемая задача выбирается из «меню» доступных ему задач или указанием пользователем имени требуемой задачи.

5. Задание исходных данных для решаемой задачи осуществляется следующим образом:

- вызов данных из локальных баз данных автоматизированных рабочих мест;
- ввод данных в режиме диалога;
- передача данных из удалённой базы данных РБНД.

6. Результаты решения задачи или ответ на запрос пользователя выводятся в виде файла, таблицы, текстового или графического документа. При необходимости результаты записываются в соответствующие базы данных.

Классификация существующих и перспективных прикладных ИТ (ПИТ) приведена на Рис. 57.

В настоящее время практически во всех ИС применяются индивидуальные ПИТ, которую эксплуатирует единственный пользователь системы, передающий полученные результаты соответствующему ЛПР.



Рис. 57

Перспективными являются групповые ПИТ, в которых над решением сложных задач работают пользователи ИС из нескольких подразделений одной организации или нескольких ИС различных организаций. Например, при «завязке» проекта нового самолета наряду с его Разработчиком такими организациями могут быть Разработчики двигателей, систем автоматического управления, бортовой электроники (авионики) и др.

Отметим, что в применении групповых технологий возникает проблема группового принятия согласованных решений различными ЛПР, имеющими собственные часто антагонистические цели.

В настоящее время в практике создания ИС различного назначения используют следующие подходы к разработке ПИТ:

1. Выбор или разработка требуемых программно-аппаратных средств, разработка алгоритмов решения задач,

программирование (кодирование) и отладка всех процедур создаваемой технологии.

2. Применение зарубежных и отечественных платформ для адаптации реализованных в них решений к задачам конкретной организации.

Отметим, что такой подход, применяемый для типовых задач в промышленности, банковской сфере и т.п., позволяет резко сократить время разработки и внедрения ПИТ. Примером такой платформы является известная с 90-х годов отечественная система «1С».

Для создания оригинальных и не имеющих мировых аналогов прикладных ИТ используется первый подход.

При таком подходе программные компоненты технологий предлагается формировать с использованием модульного принципа. Этот принцип был сформулирован в СССР в 70-х годах при разработке программного обеспечения (ПО) решения различных военно-технических задач. Согласно модульному принципу разрабатываемое ПО разбивается на отдельные элементы (программные модули), каждый из которых реализует одну или несколько функций. Для организации выполнения и взаимодействия модулей при решении конкретных задач использовался специальный управляющий (головной) модуль. Совокупность таких модулей первоначально называлась пакетом прикладных программ (ППП), а в настоящее время – комплексом программ (КП). Типовая структура КП, содержащая $(N + 1)$ модуль, представлена на Рис. 58.

Дополнительно к указанным выше функциям управляющий модуль в КП прикладных ИТ обеспечивает идентификацию пользователя, диалог с ним, взаимодействие с ОС, локальными и удаленными БД и др.

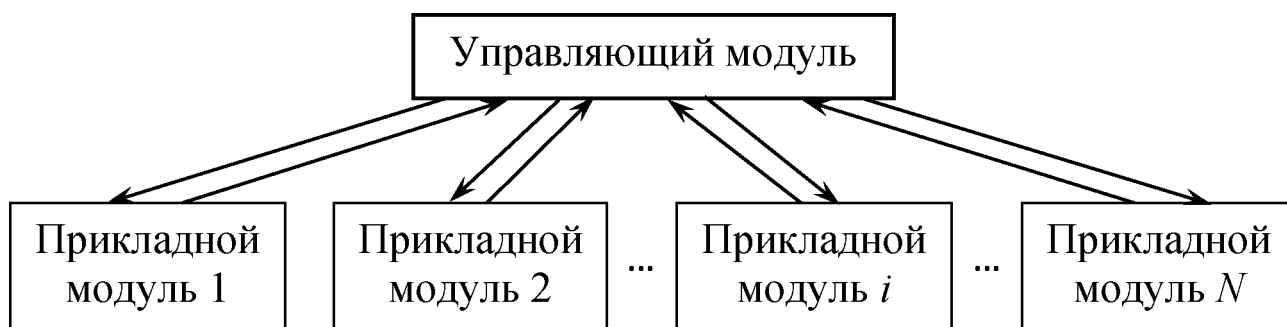


Рис. 58

Основным преимуществом модульного построения КП является незначительные затраты времени на его модификацию, которые включают в себя только замену «старого» модуля его модифицированным вариантом.

Рассмотрим различные подходы к моделированию прикладных ИТ.

Простейшая бинарная модель ПИТ, построенная по принципу «черного ящика», имеет в своей основе отношение:

$$R_{\text{пит}} \subseteq X \times Y,$$

где X – множество исходных данных для реализации задач технологии; Y – множество результатов их решения.

Если элементы множества X возникают в случайные моменты времени и на реализацию процедур ПИТ затрачивается случайное время, то для описания индивидуальных и групповых ПИТ в рамках одного подразделения предлагается использовать модели одноканальных и многоканальных систем массового обслуживания (СМО) различных типов [Вентцель Е.С.].

Для описания перспективных ПИТ, реализуемых в нескольких организациях, применяются модели сетей СМО [Кофман А., Крюон Р.].

Приведем детализированную бинарную модель прикладной ИТ.

Пусть некоторая ИС предприятия (корпорации) функционирует на основе S прикладных ИТ.

Будем считать, что для реализации s -й технологии используется определенное множество заданий Z_s , $s = (\overline{1, S})$. Примерами таких заданий являются «расчет себестоимости выпускаемой продукции», «разработка финансового плана предприятия (организации)», «расчет на прочность крыла самолета» и т.п.

Для выполнения каждого задания привлекается определенный пользователь ИС из состава множества Π_s соответствующих специалистов предприятия.

Назначение исполнителей на выполнение заданий будем описывать отношением:

$$R_{1s} \subseteq Z_s \times \Pi_s, \quad s = (\overline{1, S}).$$

Выполнение порученных заданий исполнители осуществляют с помощью множества программ (комплексов программ) P_s , $s = (\overline{1, S})$.

Это соответствие представим отношением вида:

$$R_{2s} \subseteq \Pi_s \times P_s, \quad s = (\overline{1, S}).$$

При их применении используются определенные файлы (записи) d_s базы данных B_s из состава РБНД ИС и множество Ω_s инструкций пользователям по выполнению заданий Z_s , $s = (\overline{1, S})$.

Такие распределения будем описывать следующими отношениями:

$$R_{3s} \subseteq P_s \times B_s; \quad R_{4s} \subseteq \Omega_s \times \Pi_s.$$

Результаты выполнения заданий Z_s с использованием программ множества P_s оформляются в виде совокупности электронных документов D_s , которые выдаются множеству L_s ру-

ководящего состава предприятия для анализа и принятия соответствующих решений.

Для описания завершающих этапов ПИТ будем использовать отношения вида:

$$R_{5s} \subseteq P_s \times D_s; \quad R_{6s} \subseteq D_s \times L_s; \quad R_{7s} \subseteq D_s \times B_s.$$

Последнее отношение описывает тот факт, что после соответствующих корректировок и утверждения ЛПР содержимое документов D_s записывается в соответствующие БД ИС.

В любой ИС все задания выполняются в определенные моменты ее рабочего времени. Если ввести множество $T = \{t_1, t_2, \dots, t_k\}$, описывающее рабочий календарь ИС, то план выполнения работ Z_s можно представить как:

$$R_{8s} \subseteq Z_s \times T_s.$$

Таким образом, имеем следующую бинарную модель s -й ПИТ:

$$\mathfrak{M}_s^{\text{ПИТ}} = \{S_s, \Pi_s, P_s, B_s, \Omega_s, D_s, L_s, R_{1s}, R_{2s}, \dots, R_{8s}\}.$$

Для решения задач совершенствования существующих технологий с использованием операций над отношениями $R_{1s}, R_{2s}, \dots, R_{8s}$, описанных в Разд. II, формируются вторичные отношения Q_{1s}, Q_{2s}, \dots

Например, полное описание s -й технологии можно получить из восьмиарного отношения вида:

$$Q_{5s} \subseteq Z_s \times \Omega_s \times \Pi_s \times B_s \times P_s \times D_s \times L_s \times B_s,$$

которое получается формированием следующих вторичных отношений:

$$Q_{1s} = R_{1s} \circ R_{4s}^{-1} \subseteq Z_s \times \Omega_s; \quad Q_{2s} = Q_{1s} \cup R_{4s} \subseteq Z_s \times \Omega_s \times \Pi_s;$$

$$Q_{3s} = R_{2s} \circ R_{3s} \subseteq \Pi_s \times B_s; \quad Q_{4s} = R_{3s}^{-1} \subseteq B_s \times P_s.$$

В этом случае отношение Q_{5s} получается как:

$$Q_{5s} = Q_{1s} \cup Q_{2s} \cup Q_{3s} \cup Q_{4s} \cup Q_{5s} \cup Q_{6s} \cup Q_{7s}.$$

Самостоятельная работа.

Разработать другие виды вторичных отношений и сформулировать задачи анализа ПИТ, которые могут быть решены с их использованием.

Входящие в прикладные ИТ множества заданий Z состоят из совокупности работ (задач), выполнение которых нуждается в эффективном планировании. Такое планирование предлагается осуществлять с использованием методов теории расписаний и сетевого планирования и управления (СПУ). Эти методы получили широкое распространение в 60-70 годах прошлого века, но незаслуженно забыты у нас в стране и за рубежом в настоящее время.

Рассмотрим возможное применение классической «задачи об одном станке» теории расписаний [Раскин Л.Г.].

Пусть некоторое задание Z_s включает в себя совокупность работ $z_1^{(s)}, z_2^{(s)}, \dots, z_{k_s}^{(s)}$, которые можно выполнять в произвольном порядке на одном АРМ (сервере).

Требуется выбрать оптимальную последовательность их выполнения.

Будем считать, что для каждой работы $z_i^{(s)} \in Z_s$ заданы время τ_i ее выполнения и «штраф» c_i за ее «пролеживание» в ожидании выполнения, $i = \overline{1, k_s}$.

Задача заключается в поиске оптимальной последовательности выполнения работ:

$$Z_s^{\text{опт}} = \left\{ z_{i_1}^{(s)}, z_{i_2}^{(s)}, \dots, z_{i_{k_s}}^{(s)} \right\},$$

при которой суммарная величина «штрафов» будет минимальна [Раскин Л.Г.].

Эта последовательность строится с использованием решающего правила вида:

$$\frac{\tau_{i_1}}{c_{i_1}} \geq \frac{\tau_{i_2}}{c_{i_2}} \geq \dots \geq \frac{\tau_{i_{k_s}}}{c_{i_{k_s}}},$$

с помощью которого формируется бинарное отношение предпочтения в множестве Z_s (см. Лекцию №5).

Пример.

Пусть некоторое задание Z по выполнению расчетов и формированию документов по налоговым платежам организации (корпорации) состоит в выполнении работ z_1, z_2, z_3, z_4, z_5 по пяти видам налогов.

Пусть затраты времени на расчеты, оформление и отправку документов по i -му виду налогов равны τ_i единиц времени (е.в.).

За непредставление в срок налоговым органам документов по i -му виду налогов начисляется «пени» в размере c_i денежных единиц (д.е.), $i = (\overline{1, 5})$.

Решим сформулированную выше задачу для гипотетических исходных данных, приведенных в Табл. 5.

Таблица 5

i	1	2	3	4	5
τ_i , е.в.	4	3	6	2	5
c_i , д.е.	2	1	4	3	2
τ_i/c_i	2	3	1,5	0,66	2,5

Применение решающего правила дает следующую последовательность выполнения работ:

$$Z^{\text{опт}} = \{z_1, z_5, z_1, z_3, z_4\}.$$

Для планирования выполнения заданий, реализуемых с помощью групповых ПИТ, предлагается использовать методы СПУ, по которым имеется достаточно большое число публикаций, например [Вентцель Е.С. и др.].

16. Модели управления перспективной прикладной информационной технологией интернет-рекламы

Известно, что эффективная рекламная кампания является одним из важнейших факторов продвижения на рынок произведенной продукции и услуг. В связи с бурным распространением Интернета рынок Web-рекламы стремительно растет и является весьма перспективным как для производителей товаров и услуг, так и для рекламодателей. При этом интернет-пользователи «устали» от существующей однообразной и навязчивой рекламы в виде периодически выдаваемых статичных баннеров, всплывающих окон и других объектов, мешающих им работать с выбранным сайтом.

Для привлечения внимания широкой аудитории пользователей необходимо использовать новые оригинальные информационные технологии реализации Web-рекламы. В этих технологиях предлагается применять некоторое множество S различных способов представления информации о рекламируемом товаре (услуги). Такими способами, например, являются текстовое сообщение (s_1), видеоизображение (s_2), видеоролик (s_3) и т.д.

Будем считать, что множество S содержит N элементов s_1, s_1, \dots, s_N , каждый из которых реализуется в виде прикладного программного модуля, разработанного с привлечением определенных инструментальных средств. Тогда программную реализацию Web-рекламы конкретного товара (услуги) предлагается оформить в виде комплекса программ (КП), структура которого приведена на Рис. 59. Такой КП размещается и выполняется на сайте производителя продукции или организации, предоставляющей соответствующие услуги.

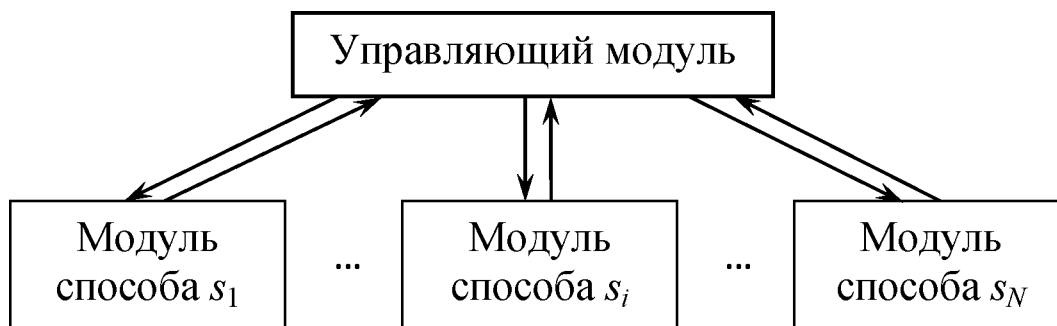


Рис. 59

При этом требуемое выполнение модулей способов s_i , $i = \overline{1, N}$ координирует управляющий модуль КП.

В этом модуле могут быть реализованы следующие методы управления Web-рекламой товара (услуги):

1. Детерминированное выполнение прикладных модулей КП.
2. Рандомизированное (случайное) выполнение этих модулей.

При детерминированном управлении Web-рекламой товара (услуги) N модулей различных способов ее представления выдаются интернет-пользователям в определенной фиксированной последовательности.

Рандомизированное управление подразумевает выдачу пользователю N видов рекламы в случайной последовательности.

Рассмотрим методы реализации этих видов управления Web-рекламой, алгоритмы которых должны быть программно реализованы в управляющем модуле КП (см. Рис. 59).

В основу работы этого модуля положим общую бинарную модель управления вида:

$$R \subseteq S \times T,$$

где $T = \{t_1, t_2, \dots, t_j, \dots, t_M\}$ — множество моментов времени выдачи пользователям рекламной информации.

Эти моменты времени будем формировать с помощью рекуррентного соотношения:

$$t_j = t_{j-1} + \Delta t, \quad j = (\overline{1, M}),$$

в котором t_j – момент времени запуска j -го модуля КП, а Δt – шаг времени запуска на выполнение модулей множества S .

При заданном значении величины Δt имеем детерминированное управление Web-рекламой. В этом случае отношение R является детерминированной моделью управления рекламой.

Если Δt является случайной величиной, распределенной по определенному закону, то имеем рандомизированный процесс выполнения модулей множества S . Тогда введенное выше отношение при множестве T , состоящем из случайных чисел, представляет собой случайное бинарное отношение \mathcal{R} , введенное в Лекции №8.

При детерминированном множестве T и заданных вероятностях p_{ij} , $i = (\overline{1, N})$, $j = (\overline{1, M})$ имеем стохастическое бинарное отношение R^c (см. Лекцию №11).

В рассматриваемых случаях работа управляющего модуля описывается вероятностными моделями.

Пусть шаг Δt является заданной детерминированной величиной. Обозначим через τ_i время выполнения модуля способа s_i , определенное в процессе его автономной отладки и тестирования. Будем использовать следующий психологический прием, который заключается в том, что в первые моменты времени выдачи рекламы пользователю выводятся «короткие» способы $s_i \in S$.

Тогда выполняемую последовательность модулей:

$$S^* = \{s_{i_1}, s_{i_2}, s_{i_3}, \dots, s_{i_N}\}$$

предлагается определять с использованием решающего правила:

$$\tau_{i_1} \leq \tau_{i_2} \leq \tau_{i_3} \leq \dots \tau_{i_N}.$$

Пример.

Пусть в КП Web-рекламы реализованы модули s_1, s_2, \dots, s_5 различных способов представления информации о рекламируемом товаре (услуге). Гипотетические значения затрат времени на их выполнение равны:

$$\tau_1 = 4 \text{ мин}; \quad \tau_2 = 3 \text{ мин}; \quad \tau_3 = 6 \text{ мин}; \quad \tau_4 = 2 \text{ мин}; \quad \tau_5 = 5 \text{ мин}.$$

Применяя приведенное выше решающее правило, получаем последовательность вида:

$$S^* = \{s_4, s_2, s_1, s_5, s_3\}.$$

Будем считать, что детерминированное множество T моментов времени работы КП определяется следующими условными значениями:

$$t_1 = 0 \text{ мин}; \quad \Delta t = 10 \text{ мин}; \quad M = 5$$

и конкретизируется как:

$$T = \{t_1, t_2, t_3, t_4, t_5\} = \{0; 10; 20; 30; 40\} \text{ мин}.$$

Тогда графовое и матричное представление модели управления Web-рекламой будут иметь, представленный на Рис. 60.

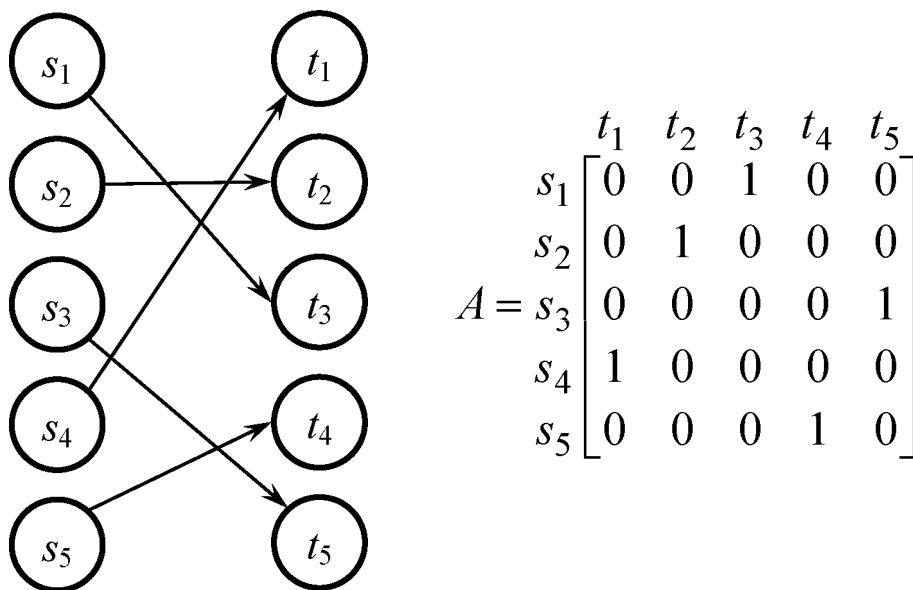


Рис. 60

Пусть множество T является множеством случайных моментов времени $t_j, j = \overline{(1, M)}$.

Будем считать, что величина Δt шага времени распределена по закону равномерной плотности (равномерному закону) в интервале значений $[80; 120]$ мин.

Функция распределения такого закона имеет вид:

$$F(\Delta t) = \frac{\Delta t - 80}{40}.$$

Согласно работам [Вентцель Е.С.; Кофман А., Крюон Р.] k -я реализация случайной величины Δt генерируется путем решения уравнения:

$$\frac{\Delta t - 80}{40} = x_k,$$

где $x_k \in (0, 1)$ – равномерное случайное число, полученное с помощью датчика случайных чисел (ДСЧ).

Решая это уравнение, получаем следующую расчетную формулу:

$$\Delta t_k = 40x_k + 80, \quad k = 1, 2, 3, \dots$$

При $t_1 = 0; M = 5; k = 1; x_k = 0,75$ и соотношении:

$$t_j^{(k)} = t_{j-1}^{(k)} + \Delta t_k, \quad j = \overline{(2, M)}$$

получаем следующую реализацию множества моментов времени запуска модулей множества S :

$$T^{(1)} = \{t_1^{(1)}, t_2^{(1)}, t_3^{(1)}, t_4^{(1)}, t_5^{(1)}\} = \{0; 109,2; 218,4 \ 327,6; 436,8\} \text{ мин.}$$

Для $k=2; x_k=0,25$ реализация множества T будет иметь вид:

$$T^{(2)} = \{0; 90; 180; 270; 360\} \text{ мин.}$$

Можно предусмотреть изменение величины Δt для получения значений $t_j^{(k)}$ по формуле:

$$t_j^{(k)} = t_{j-1}^{(k-1)} + \Delta t_k, \quad k = 1, 2, 3, \dots, \quad j = 2, 3, \dots$$

Тогда первые три элемента множества T при $t_1 = 0$; $x_1 = 0,73$; $x_2 = 0,25$ и величине шага $\Delta t_1 = 109,2$; $\Delta t_2 = 90$ будут равны:

$$t_1 = 0; \quad t_2 = 109,2 \text{ мин}; \quad t_3 = 199,2 \text{ мин}.$$

Рассмотрим реализацию рандомизированного управления Web-рекламой, в основу которого полагается методология методов Монте-Карло. При этом будем использовать метод статистического моделирования процесса наступления случайных событий [Вентцель Е.С.].

Будем считать, что использование в каждый конкретный момент времени s_i -го способа предоставления рекламной информации является случайным событием B_i с известной вероятностью p_i его наступления, $i = 1, 2, 3, \dots, N$. Эти вероятности должны удовлетворять условию:

$$p_1 + p_2 + \dots + p_N = 1, \quad (94)$$

которое означает, что хотя бы одно из случайных событий B_1, B_2, \dots, B_N должно произойти.

Метод Монте-Карло позволяет моделировать наступление полной группы случайных событий B_1, B_2, \dots, B_N с использованием датчика равномерно распределенных в интервале значений от 0 до 1 случайных чисел [Вентцель Е.С.].

Для этих целей сформируем вспомогательные вероятности:

$$q_i = p_1 + p_2 + \dots + p_i, \quad i = 1, 2, 3, \dots, N. \quad (95)$$

Из сопоставления этого выражения с формулой (94) следует, что при $i = N$ имеем:

$$q_N = 1. \quad (96)$$

Разобьем числовой интервал значений от 0 до 1 точками $q_1, q_2, \dots, q_{N-1}, q_N$ на отрезки так, как это показано на Рис. 61.

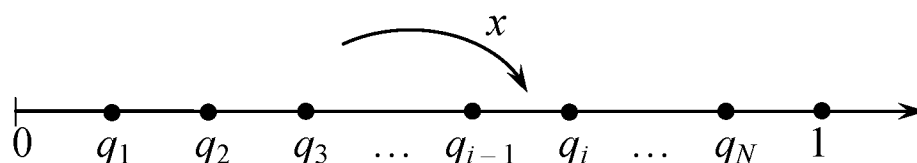


Рис. 61

В предлагаемом методе [Вентцель Е.С.] утверждается, что если сгенерированное в некоторый момент времени t датчиком случайных чисел (ДСЧ) значение $x_t \in (0, 1)$ удовлетворяет неравенству:

$$q_{i-1} < x_t < q_i, \quad (97)$$

то считается, что в этот момент времени наступило событие B_i , $i \in \{1, 2, 3, \dots, N\}$.

Входящие в это условие значения вероятностей q_i вычисляются по формулам (95) и (96). При этом считается, что:

$$q_0 = 0.$$

Простейший алгоритм вышеизложенного метода, который предлагается использовать в управляющем модуле перспективных КП интернет-рекламы, включает в себя следующие этапы:

1°. Ввод значения применяемого числа N способов представления информации о товаре (услуге).

2°. Вычисление значений вероятностей применения каждого способа:

$$p_i = \frac{1}{N}, \quad i = 1, 2, 3, \dots, N. \quad (98)$$

3°. Выбор шага изменения времени Δt смены способов предоставления рекламой информации.

4°. Генерация числа $x_t \in (0, 1)$ в каждый момент времени:

$$t = t + \Delta t$$

с помощью ДСЧ.

5°. Проверка условия вида (97), которое с учетом формул (95) и (98) представляется как:

$$\frac{(i-1)}{N} < x_t < \frac{i}{N}, \text{ для всех } i = 1, 2, 3, \dots, N. \quad (99)$$

6°. Фиксация значения i , для которого неравенство (99) выполнилось.

7°. Запуск модуля s_i , $i \in \overline{(1, N)}$ в момент времени t .

8°. Переход к п. 4°.

Пример.

Будем считать, что в Web-рекламе некоторой продукции (услуги) было использовано $N = 3$ следующих способа ее отображения:

s_1 – Статическое изображение;

s_2 – Выплывающее изображение;

s_3 – Аудио-видеоизображение.

Вероятности их применения, вычисленные по формуле (98), будут равны:

$$p_1 = \frac{1}{3}, \quad p_2 = \frac{1}{3}, \quad p_3 = \frac{1}{3}.$$

Тогда вероятности (95) примут следующие значения:

$$q_0 = 0; \quad q_1 = \frac{1}{3} = 0,333; \quad q_2 = \frac{1}{3} = 0,666; \quad q_3 = 1.$$

Неравенства (97) или аналогичные им условия (99) запишутся как:

$$\begin{aligned} 0 < x_t < 0,333; \\ 0,333 < x_t < 0,666; \\ 0,666 < x_t < 1. \end{aligned} \quad (100)$$

Будем считать, что шаг изменения способа отображения информации равен $\Delta t = 5$ мин.

В Табл. 4 приведена динамика изменения способов отображения рекламы в течение одного часа, использующая анализ выполнения неравенств (100).

Таблица 4

Время t , мин.	Случайное число x_t	Способ отображения рекламы
0	0,73	s_3
5	0,25	s_1
10	0,45	s_2
15	0,58	s_2
20	0,20	s_1
25	0,54	s_2
30	0,01	s_1
35	0,21	s_1
40	0,51	s_2
45	0,96	s_3
50	0,35	s_2
55	0,61	s_2
60	0,42	s_2

В общем случае матрица A стохастического отношения $R^c \subseteq S \times T$ будет иметь следующий вид:

$$A = \begin{matrix} & t_1 & t_2 & \dots & t_M \\ \begin{matrix} s_1 \\ s_2 \\ \dots \\ s_N \end{matrix} & \begin{bmatrix} p_1 & p_1 & p_1 & p_1 \\ p_2 & p_2 & p_2 & p_2 \\ \dots & \dots & \dots & \dots \\ p_N & p_N & p_N & p_N \end{bmatrix} \end{matrix}.$$

Эта матрица используется управляющим модулем при генерации реализаций рандомизированного выполнения прикладных модулей КП Web-рекламы.

17. Модели распределенного банка данных прикладных информационных технологий

Для решения задач анализа и синтеза прикладных информационных технологий (ПИТ) кроме использования математических моделей информационных систем (ИС) необходимо использовать частные модели их компонент, таких как:

1. Модели информационного обеспечения работы ИС.
2. Модели программных компонент ИС.

Отметим, что в существующей литературе, отсутствуют общепринятые модели этих и других компонент ИС. Исключение составляют модели комплексов технических средств, которые формализуются в основном с привлечением математического аппарата теории массового обслуживания и теории графов.

Следует отметить, что за счет «больших размеров» современных ИС и невозможности чёткой регламентации работы всех элементов, модели их компонент должны в основном рассматриваться как вероятностные модели.

В этой и последующей Лекциях будут рассмотрены вопросы моделирования этих компонент ПИТ.

Современное информационное обеспечение работы крупных организаций (корпораций) осуществляется с помощью их банков данных (БнД).

Существует два подхода к использованию БнД в ИС:

1. Применение единого централизованного банка организаций (корпораций). Таким БнД, получившим широкое распространение у нас в стране является БнД «Oracle».
2. Использование в составе ИС распределенного БнД (РБнД), состоящего из совокупности локальных БнД подразделений организации (корпорации) и специальной СУРБД.

Основными недостатками первого подхода являются:

- большой (сверхбольшой) объем хранимых данных требует значительного количества аппаратных средств (серверов) и отрицательно сказывается на затратах времени по сопровождению БНД и поиску требуемой для выполнения ПИТ информации;

- малая надежность функционирования ИС за счет возможных сбоев, отказов аппаратуры и СУБД, а также успешных информационных атак на серверы БНД.

При оценке функциональной надежности ИС можно воспользоваться методами и статистикой, изложенными в монографии [Зиновьев П.А., Моисеев В.С., Мейко А.В.].

Модели работы пользователей ПИТ с единым БНД ИС представлены в Лекции №13.

Рассмотрим модели работы РБНД для выполнения индивидуальных и групповых ПИТ крупных корпоративных ИС (КИС), представленных на Рис. 57.

Будем считать, что в каждом j -м подразделении организации (корпорации) создается локальный БНД (ЛБНД) \mathfrak{B}_j , с которым работает с использованием соответствующих АРМ множество Π_j сотрудников этого подразделения, $j = (\overline{1, n})$. Эти банки могут взаимодействовать между собой путем обмена информацией с помощью системы связи ИС (КИС).

Для построения вероятностной функциональной модели РБНД будем использовать развитие теории линейных стохастических сетей систем массового обслуживания (СМО), изложенной в Лекции №14. В этой модели за счет случайности потоков запросов пользователей к ЛБНД, случайного времени их обработки и случайного характера взаимодействия между

пользователями индивидуальных и групповых ПИТ системы в качестве СМО выступают ее ЛБНД.

Общая структура РБНД, которая идентична структуре организации (корпорации), приведена на Рис. 62.

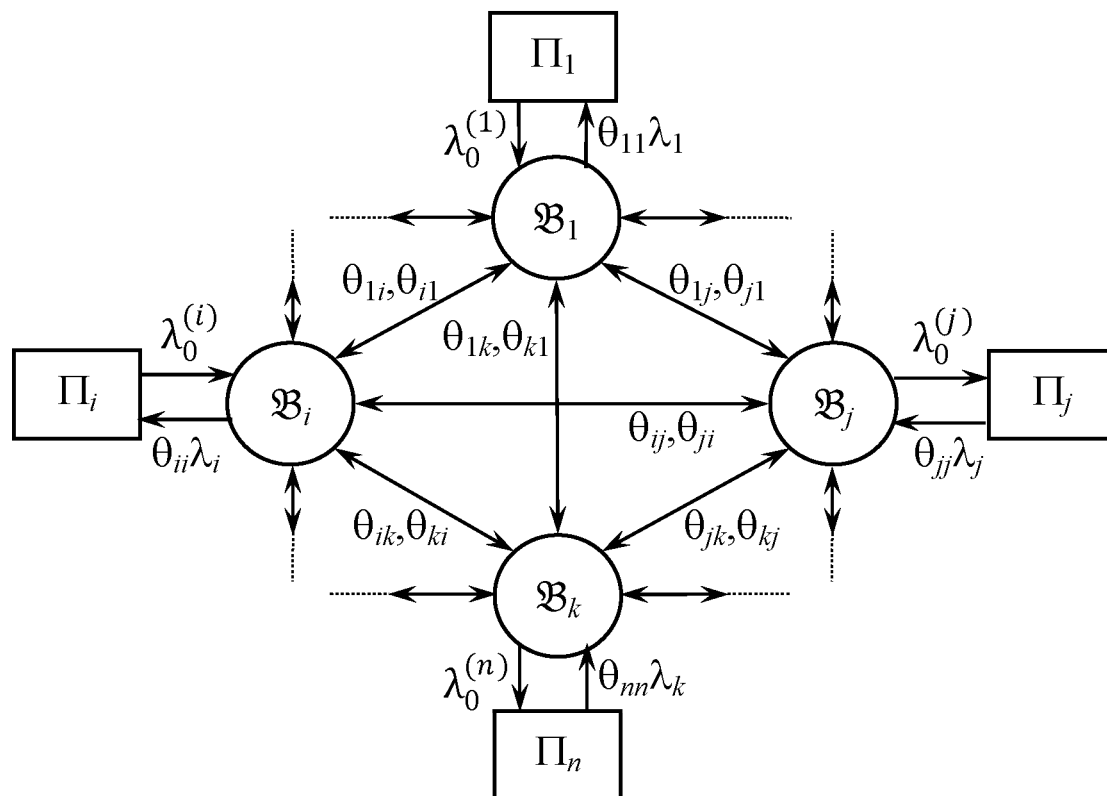


Рис. 62

На этом рисунке величины $\lambda_0^{(j)}$ означают интенсивность запросов к РБНД множества пользователей Π_j , $j = (\overline{1, n})$. Для простоты представления схемы взаимодействия ЛБНД вероятности θ_{ij} и θ_{ji} указаны на одной дуге графа передач сети, описывающей взаимосвязь банков \mathfrak{B}_i и \mathfrak{B}_j , $i, j = (\overline{1, n})$.

Будем считать, что граф передач запросов сети РБНД является полным графом [Харари Ф.], в котором каждая вершина связана дугами со всеми остальными вершинами.

На Рис. 63 приведен частный случай структуры РБНД для значения $n = 3$.

При построении модели РБнД будем учитывать следующие особенности решаемой задачи:

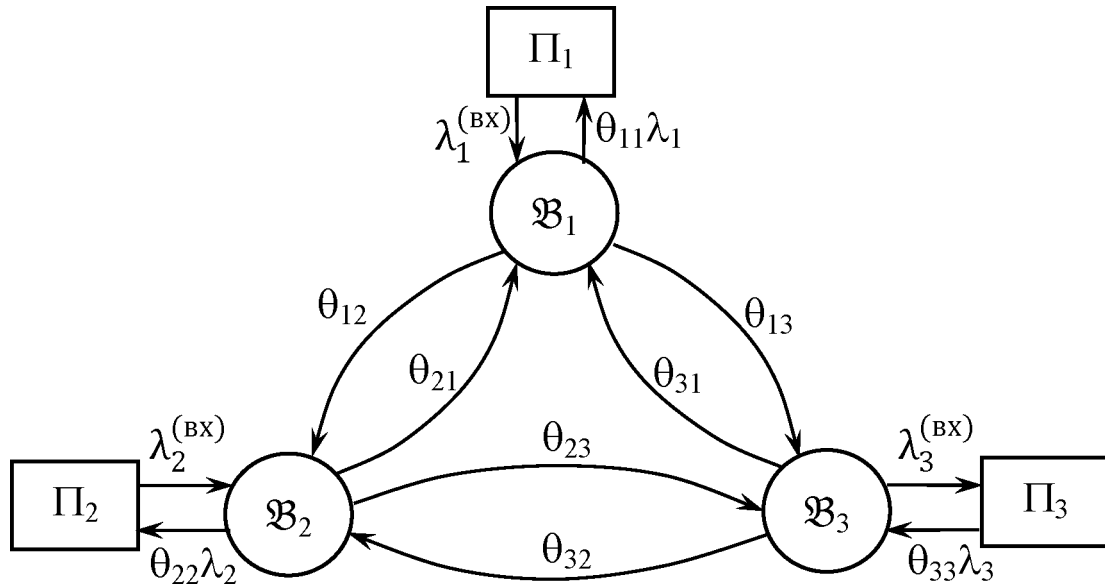


Рис. 63

1) линейная стохастическая сеть ЛБнД имеет n источников заявок на обслуживание (запросов к РБнД),

2) вероятности θ_{ij} матрицы T передач запросов пользователей множества Π_i означают вероятность их удовлетворения информацией, хранимой в базах данных банка \mathfrak{B}_i , $i = (\overline{1, n})$,

3) интенсивности потоков запросов λ_i означают интенсивности обращения баз данных \mathfrak{B}_j к другим БнД сети,

4) интенсивности неудовлетворенных ЛБнД обращений пользователей, которые «уходят» для обслуживания в сеть, определяются величиной $\lambda_j^{(BX)}(1 - \theta_{jj})$, где $\lambda_j^{(BX)}$ – интенсивность входного потока запросов от пользователей множества Π_j , $j = (\overline{1, n})$.

5) интенсивности выходных потоков запросов вычисляются по формуле вида:

$$\lambda_j^{(ВЫХ)} = \theta_{jj} \lambda_j, \quad j = (\overline{1, n}). \quad (101)$$

Определим матрицу передач запросов между ЛБНД сети следующим образом:

$$T = \begin{matrix} & \mathfrak{B}_1 & \mathfrak{B}_2 & \dots & \mathfrak{B}_n \\ \mathfrak{B}_1 & \theta_{11} & \theta_{12} & \dots & \theta_{1n} \\ \mathfrak{B}_2 & \theta_{21} & \theta_{22} & \dots & \theta_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \mathfrak{B}_n & \theta_{n1} & \theta_{n2} & \dots & \theta_{nn} \end{matrix} \quad (102)$$

Значения элементов этой матрицы $\theta_{ij} \in [0; 1]$, $i, j = (\overline{1, n})$ должны удовлетворять условиям нормировки вида:

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

Система линейных неоднородных алгебраических уравнений для определения интенсивностей λ_j при заданных значениях $\lambda_j^{(\text{BX})}$ и θ_{ij} , $i, j = (\overline{1, n})$ будет иметь следующий вид:

$$\begin{aligned} \lambda_1 &= \lambda_1^{(\text{BX})}(1 - \theta_{11}) + \sum_{i=2}^n \lambda_i \theta_{i1}; \\ &\dots\dots\dots \\ \lambda_j &= \sum_{i=1}^{j-1} \lambda_i \theta_{ij} + \lambda_j^{(\text{BX})}(1 - \theta_{jj}) + \sum_{i=j+1}^n \lambda_i \theta_{ij}, \quad j = (\overline{2, n-1}); \quad (103) \\ &\dots\dots\dots \\ \lambda_n &= \sum_{i=1}^{n-1} \lambda_i \theta_{in} + \lambda_n^{(\text{BX})}(1 - \theta_{nn}). \end{aligned}$$

Эта система имеет ранг, равный n и всегда имеет решение, которое может быть получено с использованием известных численных методов [Вержбицкий В.М.].

Таким образом, простейшая вероятностная модель функционирования РБНД описывается выражениями (101)-(103).

Проведем исследование полученной модели.

Если считать, что все запросы пользователей каждого подразделения организации (корпорации) полностью удовлетворяются его ЛБнД, то вероятности θ_{jj} должны быть тождественно равны единице. Но из приведенных выше условий нормировки следует, что остальные вероятности принимают следующие значения:

$$\theta_{ij} = 0, \quad i, j = (\overline{1, n}), \quad i \neq j.$$

Тогда из выражений (103) следует, что:

$$\lambda_j = 0; \quad j = (\overline{1, n}).$$

Эти неравенства означают отсутствие потоков в сети РБнД.

Таким образом, условия применения предлагаемой модели имеют вид:

$$0 < \theta_{jj} < 1, \quad j = (\overline{1, n}).$$

При $\theta_{jj} = 0$ предполагается, что пользователи множества Π_j осуществляют только сбор и ввод информации в ЛБнД \mathfrak{B}_j , $j = (\overline{1, n})$, с которыми будут работать другие пользователи РБнД.

Для решения задач анализа РБнД с помощью ее модели определим количество запросов, проходящих через каждый ЛБнД:

$$\alpha_j = \frac{\lambda_j^{(0)}}{\lambda_j^{(BX)}}, \quad j = (\overline{1, n}), \quad (104)$$

где $\lambda_j^{(0)}$ – решения системы уравнений (103).

Номера наиболее и наименее нагруженных запросами ЛБнД \mathfrak{B}_j определяются как:

$$i_{\max} = \arg \max_{j=(\overline{1, n})} \{\alpha_j\}; \quad i_{\min} = \arg \min_{j=(\overline{1, n})} \{\alpha_j\}. \quad (105)$$

Среднее время обработки запросов в РБнД вычислим по известной формуле [Кузин Л.Т.; Кофман А., Крюон Р.]:

$$\tau = \sum_{j=1}^n \alpha_j \tau_j, \quad (106)$$

где τ_j – среднее время пребывания запроса в ЛБнД \mathfrak{B}_j , $j = (\overline{1, n})$.

Для показательных сетей СМО [Кузин Л.Т.; Кофман А., Крюон Р.] должны выполняться условия наличия установившегося режима РБнД:

$$\rho_j = \frac{\lambda_j^{(0)}}{\mu_j^{(\text{вх})}} < 1, \quad j = (\overline{1, n}),$$

где $\mu_j = \tau_j^{-1}$ – интенсивность обслуживания потока запросов в ЛБнД \mathfrak{B}_j , $j = (\overline{1, n})$.

Пример.

Проведем расчеты для структуры РБнД, приведенного на Рис. 63.

Матрица (102) записывается как:

$$T = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix}.$$

Система уравнений (103) для этой матрицы примет вид:

$$\begin{aligned} \lambda_1 &= \lambda_1^{(\text{вх})}(1 - \theta_{11}) + \lambda_2 \theta_{21} + \lambda_3 \theta_{31}; \\ \lambda_2 &= \lambda_1 \theta_{12} + \lambda_2^{(\text{вх})}(1 - \theta_{22}) + \lambda_3 \theta_{32}; \\ \lambda_3 &= \lambda_1 \theta_{13} + \lambda_2 \theta_{23} + \lambda_3^{(\text{вх})}(1 - \theta_{33}). \end{aligned}$$

Будем считать, что 60% запросов пользователей Π_1 удовлетворяются информацией, хранящейся в ЛБнД \mathfrak{B}_1 . Для остальных пользователей РБнД предполагаем, что такие проценты будут соответственно равны 40% (база \mathfrak{B}_2) и 50% (ЛБнД \mathfrak{B}_3).

С учетом этого элементы матрицы T примут следующие значения:

$$\begin{aligned}\theta_{11} &= 0,6; \quad \theta_{12} = 0,3; \quad \theta_{13} = 0,1; \\ \theta_{12} &= 0,35; \quad \theta_{22} = 0,4; \quad \theta_{23} = 0,25; \\ \theta_{31} &= 0,25; \quad \theta_{32} = 0,25; \quad \theta_{33} = 0,5.\end{aligned}$$

Будем считать, что входные интенсивности запросов пользователей и интенсивности их обработки в ЛБНД сети соответственно равны:

$$\begin{aligned}\lambda_1^{(\text{вх})} &= \frac{1}{5} \text{ с}^{-1}; \quad \lambda_2^{(\text{вх})} = \frac{1}{8} \text{ с}^{-1}; \quad \lambda_3^{(\text{вх})} = \frac{1}{10} \text{ с}^{-1}; \\ \mu_1 &= \frac{1}{4} \text{ с}^{-1}; \quad \mu_2 = \frac{1}{5} \text{ с}^{-1}; \quad \mu_3 = \frac{1}{3} \text{ с}^{-1}.\end{aligned}$$

Решая приведенную выше систему уравнений, имеем, что:

$$\lambda_1^{(0)} = 0,157 \text{ с}^{-1}; \quad \lambda_2^{(0)} = 0,148 \text{ с}^{-1}; \quad \lambda_3^{(0)} = 0,103 \text{ с}^{-1}.$$

Интенсивности выходных потоков обслуженных запросов пользователей, вычисленные при $\lambda_j = \lambda_j^{(0)}$ по формулам (101), будут равны:

$$\lambda_1^{(\text{вых})} = 0,0942 \text{ с}^{-1}; \quad \lambda_2^{(\text{вых})} = 0,0592 \text{ с}^{-1}; \quad \lambda_3^{(\text{вых})} = 0,0515 \text{ с}^{-1}.$$

Среднее число проходов запросов через ЛБНД определим по формуле (104):

$$\alpha_1 = 0,785; \quad \alpha_2 = 1,184; \quad \alpha_3 = 1,030.$$

Из этих значений и соотношений (105) следует, что наиболее нагруженным является банк \mathfrak{B}_2 , а наименее нагруженным — \mathfrak{B}_1 . Объяснением этому являются значения вероятностей θ_{22} и θ_{11} удовлетворения запросов пользователей соответствующими ЛБНД.

Используя выражение (106), определим среднее время выполнения запросов к РБНД:

$$\tau = \tau_1 + \tau_2 + \tau_3 = 0,785 \cdot 4 + 1,184 \cdot 5 + 1,030 \cdot 3 = 12,15 \text{ с}.$$

Будем считать, что в ЛБнД \mathfrak{B}_1 собираются данные по отказам и неисправностям производственного оборудования, установленного в первом ($j = 1$) подразделении оргструктуры рассматриваемого гипотетического предприятия. Эти данные используют в своей деятельности пользователи двух других подразделений предприятия.

Таким образом, элементы первой строки матрицы T будут иметь следующий вид:

$$\theta_{11} = 0; \quad \theta_{12} = 0,5; \quad \theta_{13} = 0,5.$$

Для этих значений и приведенных выше значений θ_{ij} и $\lambda_i^{(\text{вх})}$, $i, j = (\overline{1, 3})$ решения системы уравнений (103) будут равны:

$$\lambda_1^{(0)} = 0,418 \text{ с}^{-1}; \quad \lambda_2^{(0)} = 0,372 \text{ с}^{-1}; \quad \lambda_3^{(0)} = 0,352 \text{ с}^{-1}.$$

Выходные потоки РБнД имеют следующие значения:

$$\lambda_1^{(\text{вых})} = 0 \text{ с}^{-1}; \quad \lambda_2^{(\text{вых})} = 0,149 \text{ с}^{-1}; \quad \lambda_3^{(\text{вых})} = 0,176 \text{ с}^{-1}.$$

Среднее количество прохождений запросов через ЛБнД будет равно:

$$\alpha_1 = 2,09; \quad \alpha_2 = 3,02; \quad \alpha_3 = 3,52.$$

Отсюда следует, что максимальную нагрузку имеет банк \mathfrak{B}_3 , а наименее загруженным будет банк \mathfrak{B}_1 .

Среднее время выполнения запросов к РБнД в рассматриваемом случае вычисляется как:

$$\tau = \tau_1 + \tau_2 + \tau_3 = 2,09 \cdot 4 + 3,02 \cdot 5 + 3,52 \cdot 3 = 34,02 \text{ с}.$$

Из сравнения с предыдущим результатом видно, что время реакции РБнД выросло почти в 3 раза (*почему?*).

Предложенная выше модель функционирования РБнД относится к классу статических вероятностных моделей, которая описывает ее работу в установившемся режиме.

Приведем простейшую динамическую вероятностную модель функционирования РБнД.

Из теории дифференциальных уравнений известно, что для автономной системы:

$$\frac{dx_i}{dt} = \varphi_i(x_1, x_2, \dots, x_n), \quad i = (\overline{1, n}),$$

для которой выполняются условия:

$$\frac{\partial \varphi_i}{\partial x_i} < 0, \quad i = (\overline{1, n}),$$

при $t \rightarrow \infty$ решения $x_i(t)$ асимптотически приближаются к значениям $x_i^{(0)}$, $i = (\overline{1, n})$, которые являются корнями системы алгебраических уравнений:

$$\varphi_i(x_1, x_2, \dots, x_n) = 0, \quad i = (\overline{1, n})$$

Перепишем выражения (103) в виде следующей системы дифференциальных уравнений:

$$\dot{\lambda}_1 = -\lambda_1 + \lambda_1^{(\text{BX})}(1 - \theta_{11}) + \sum_{i=2}^n \lambda_i \theta_{i1};$$

.....

$$\dot{\lambda}_j = -\lambda_j + \sum_{i=1}^{j-1} \lambda_i \theta_{ij} + \lambda_j^{(\text{BX})}(1 - \theta_{jj}) + \sum_{i=j+1}^n \lambda_i \theta_{ij}, \quad j = (\overline{2, n-1});$$

.....

$$\dot{\lambda}_n = -\lambda_n + \sum_{i=1}^{n-1} \lambda_i \theta_{in} + \lambda_n^{(\text{BX})}(1 - \theta_{nn}), \quad t \in [t_0, t_k],$$

где $\dot{\lambda}_j$ – производные интенсивностей $\lambda_j(t)$ входных потоков запросов к ЛБНД \mathfrak{B}_j от времени t ; t_0 – момент времени начала эксплуатации РБНД.

Эта система интегрируется при начальных условиях:

$$\lambda_j(t_0) = 0, \quad j = (\overline{1, n}).$$

Решения $\lambda_j(t)$, $t \in [t_0, t_k]$ приведенной выше задачи Коши позволяют определить динамику показателей(104)-(106)

функционирования РБнД. Кроме этого модель позволяет определить момент времени $T > t_0$ выхода РБнД в установившийся режим функционирования и предельные значения интенсивностей $\lambda_j^{(0)} = \lambda_j(T)$, $j = (\overline{1, n})$.

Для представленных выше примеров применения статической модели РБнД использование динамической модели при $t \in [0; 60]$ с. дает величину $T \approx 12$ с. в первом из них и $T \approx 18$ с. во втором примере. В эти моменты времени интенсивности $\lambda_1(t)$, $\lambda_2(t)$, $\lambda_3(t)$, монотонно возрастаая, принимают значения $\lambda_j^{(0)}$, $j = (\overline{1, 3})$, полученные с использованием статической модели (103).

Если ввести в рассмотрение множество ЛБнД, реализованных в составе РБнД КИС:

$$\mathfrak{B} = \{\mathfrak{B}_1, \mathfrak{B}_2, \dots, \mathfrak{B}_n\},$$

то можно описать их взаимодействие в виде бинарной статической модели:

$$R^c \subseteq \mathfrak{B} \times \mathfrak{B},$$

где R^c – стохастическое отношение, введенное в Лекции №11.

Это отношение задается матрицей передач запросов, заданной выражением (102).

Использование метода построения реализаций отношения R^c , представленного в указанной лекции, позволяет осуществлять имитационное моделирование взаимодействия ЛБнД.

18. Модели комплексов программ прикладных информационных технологий

Рассмотрим вопросы моделирования программного обеспечения прикладных информационных технологий (ПИТ), которые являются главными элементами ИС.

Программное обеспечение (ПО) ИС можно рассматривать как совокупность взаимодействующих во времени и пространстве комплексов системных и прикладных программ.

В СССР для ПО автоматизированных систем различного назначения, изготовляемого на основе промышленного (индустриального) подхода, который подразумевал широкое использование на практике разнообразных систем автоматизации программирования (САП), было введено следующее понятие программного изделия (продукта).

Программное изделие (продукт) – программа на машинном носителе, сопровождаемая программной и эксплуатационной документацией. В практике их использования они носили названия пакетов программ, комплексов программ, программных комплексов, программных систем и т.п.

Для отдельных компонент программного изделия использовалось понятие программных модулей.

В общем случае работу любого программного изделия (ПИ) можно представить в виде схемы, представленной на Рис. 64.

Из этой схемы следует, что программное изделие – это совокупность взаимодействующих в процессе решения задач программных модулей, получающих из различных источников входные данные в виде множеств X и выдающих потребителям результаты работы в виде множеств Y .

Понятие программного модуля (ПМ) является развитием понятия «подпрограмма», широко применяемого в отечественном программировании 60-70-х годов прошлого века. Под ним понималась часть программы решения сложной задачи, которая выполняла одну конкретную функцию (вычисление тригонометрических функций, определенного интеграла и

т.п.). К ранним системам программирования ЭВМ обычно прилагалась библиотека стандартных подпрограмм с инструкциями по их применению.

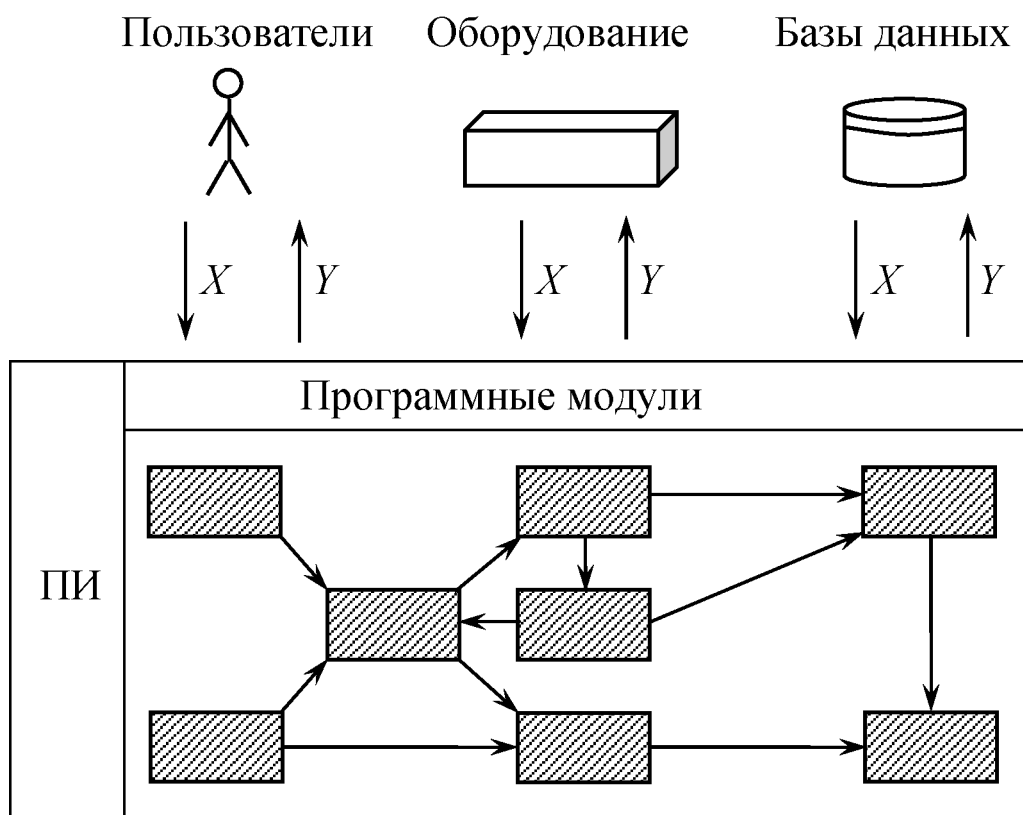


Рис. 64

В 70-80-е г.г. был сформулирован принцип модульного программирования сложных задач, в котором под ПМ понимался отдельный элемент «большой» программы (ПИ), выполняющий одну или несколько ее функций. При этом предполагалось, что модуль достаточно просто должен «извлекаться» из текста «большой» программы (ПИ), модифицироваться и вновь размещаться в ее новой версии.

В настоящее время программные модули разрабатываются с использованием универсальных (Си, Ява и др.) и проблемно-ориентированных (1С и др.) языков программирования.

Отметим еще одну особенность модульного принципа построения ПИ, которая состоит в том, что его модули должны иметь возможность автономной разработки, отладки и тести-

рования различными программистами с последующей комплексной отладкой и тестированием в составе ПИ.

Для моделирования ПМ обычно используют графовые модели программ, предложенные А.П. Ершовым в его монографии по теоретическому программированию.

Графом управления программного модуля будем называть ориентированный граф $G(O, U)$, вершинами которого является множество O операторов программы, а множество дуг U описывает всевозможные передачи управления в процессе выполнения модуля. Пример графа $G(O, U)$ для простейшей программы приведен на Рис. 65.

Текст программы

```

 $o_1$ . begin
 $o_2$ . read ( $a, b, c$ )
 $o_3$ . if ( $a < 0$ ) go to  $o_8$ 
 $o_4$ .  $y := a + b + c$ 
 $o_5$ . write  $a, b, c, y$ 
 $o_6$ .  $a := a + 1$ 
 $o_7$ . if ( $a \leq 10$ ) go to  $o_4$ 
 $o_8$ . end

```

Граф $G(O, U)$

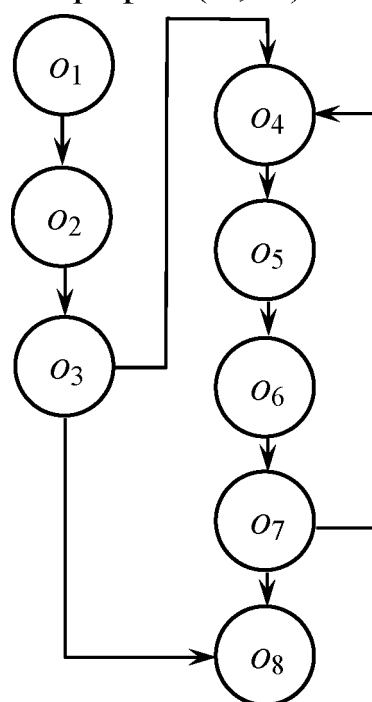


Рис. 65

Заметим, что графовые модели программ также можно рассматривать как их бинарные модели, используя определение ориентированного графа в виде отношения [Кристофидес Н.]:

$$G \subseteq O \times O, \quad (107)$$

где $O = \{o_1, o_2, \dots, o_n\}$ – подмножество операторов конкретного языка программирования, применяемого при разработке ПМ.

Графы $G(O, U)$ могут быть использованы для решения всевозможных задач анализа программ. Различные алгоритмы анализа графов с многочисленными примерами, которые можно использовать при анализе программ, приведены в работе [Кристофидес Н.]. В частности, с их помощью можно решать задачу анализа корректности программ, т.е. наличия в тексте программ «висячих» и «тупиковых» операторов, а также тестирования сложных программ.

Отметим, что в СССР были созданы программные системы, позволяющие автоматически формировать и анализировать такие графы для ПИ, содержащих десятки тысяч операторов.

Согласно Лекции №5, отношение (107) описывается матрицей $A = [a_{ij}]_{n \times n}$ и соответствующим двудольным графом G .

Для примера, представленного на Рис. 65, такие описания имеют вид, показанный на Рис. 66.

Статическая детерминированная модель ПМ (107) может быть преобразована в его стохастическую G^c и динамическую G^t бинарные модели. В таких моделях для элементов $o_i \in O$, осуществляющих условные передачи управления другим операторам ПМ, в матрицах A элементы $a_{ij} \in \{0; 1\}$ заменяются на элементы p_{ij} и кусочно-постоянные функции $a_{ij}(t)$, $t \in [0, T]$.

В приведенном примере для модели G^c такими элементами являются:

$$a_{34} = p_{34}; \quad a_{38} = p_{38};$$

$$a_{74} = p_{74}; \quad a_{78} = p_{78}.$$

При этом, согласно Лекции №11, должны выполняться условия:

$$p_{34} + p_{38} = 1; \quad p_{74} + p_{78} = 1.$$

$$A = \begin{matrix} & \begin{matrix} o_1 & o_2 & o_3 & o_4 & o_5 & o_6 & o_7 & o_8 \end{matrix} \\ \begin{matrix} o_1 \\ o_2 \\ o_3 \\ o_4 \\ o_5 \\ o_6 \\ o_7 \\ o_8 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

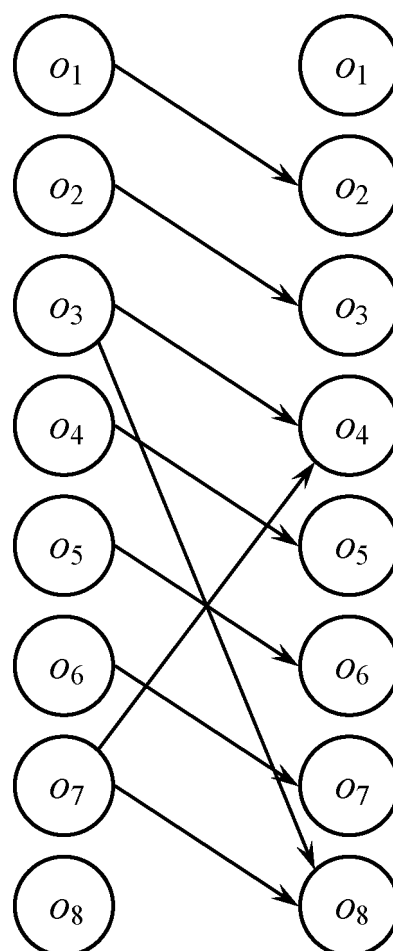


Рис. 66

При динамическом моделировании работы этой программы на заданном интервале времени в матрице $A(t)$, согласно Лекции №10, используются элементы $a_{34}(t)$, $a_{38}(t)$, $a_{74}(t)$, $a_{78}(t)$ в виде кусочно-постоянных функций.

Применение этих новых видов бинарных отношений позволяет для решения задач анализа ПМ использовать имитационное моделирование их функционирования, описанное в отмеченных выше Лекциях №10 и №11.

Основным программным компонентом ПИТ является комплекс программ (КП), общая структура которого приведена на Рис. 58 (Лекция №15).

Обозначим через $M = \{m_1, m_2, \dots, m_n\}$ множество прикладных модулей КП и через m_0 его управляющий модуль. При решении конкретных задач (выполнение работ, входящих в

ПИТ) этот модуль запускает в определенной последовательности модули множества M (см. Рис. 64).

Будем описывать взаимосвязи модулей $m_i \in M$ при эксплуатации КП ориентированным графом $\Gamma(M, U)$. Как и выше эти связи можно представить бинарным отношением вида:

$$R_1 \subseteq M \times M, \quad (108)$$

которое используется при функционировании управляющего модуля КП.

При поступлении определенных совокупностей входных данных $X_i \subseteq X$ этот модуль запускает на выполнение упорядоченную последовательность прикладных модулей $\{m_{i_1}, m_{i_2}, \dots, m_{i_{s_r}}\}$. Будем считать, что параметр i_1 является номером первого выполняемого модуля множества M при решении с использованием КП r -й задачи ПИТ.

Для формирования полного множества таких последовательностей сформируем с помощью графа $\Gamma(M, U)$ методами теории графов полное множество L маршрутов (путей). Каждый маршрут $l_r \in L$ представляет собой приведенную выше последовательность прикладных модулей множества M . Состав маршрутов выполнения этих модулей описывается бинарным отношением:

$$R_2 \subseteq L \times M.$$

Для конкретизации работы КП с множествами входных X и выходных Y данных рассматриваемой ПИТ введем в рассмотрение следующие отношения:

$$R_3 \subseteq X \times M; \quad R_4 \subseteq M \times Y.$$

Тогда статическая детерминированная модель КП некоторой ПИТ будет иметь следующий вид:

$$\mathfrak{M}_{\text{кп}} = \{M, X, Y, L, R_1, R_2, R_3, R_4\}.$$

Для решения задач анализа КП на основе этой модели можно сформировать ряд вторичных отношений.

Например, тернарное отношение, описывающее связь входных данных, модулей и результатов их работы, получается как:

$$Q_1 = R_3 \cup R_4 \subseteq X \times M \times Y.$$

Связь данных, необходимых для выполнения определенных маршрутов, можно представить отношением вида:

$$Q_2 = R_3 \circ R_2^{-1} \subseteq X \times Y.$$

Результаты работы КП при реализации маршрутов выполнения его прикладных модулей представляются как:

$$Q_3 = R_2 \circ R_4 \subseteq L \times Y.$$

Значительное число свойств рассматриваемого КП может быть получено при анализе графа $\Gamma(M, U)$ известными методами и алгоритмами теории графов [Кристофидес Н.; Харари Ф. и др].

Приведем метод определения минимальных t_{\min} и максимальных t_{\max} затрат времени выполнения прикладных модулей множества M при решении с помощью рассматриваемого КП всех задач ПИТ.

Введем в рассмотрение ориентированный граф $\Gamma(M, U)$ с множеством вершин $M = \{m_1, m_2, \dots, m_n\}$ и множеством U ребер (дуг), описывающих всевозможные передачи управления между модулями m_i и m_j , $i, j \in (\overline{1, n})$ в процессах выполнения функционала КП.

Будем считать, что при автономной разработке прикладных модулей КП определены затраты времени τ_i на выполнение каждого прикладного модуля m_i , $i \in (\overline{1, n})$. Затраты времени функционирования управляющего модуля КП предпола-

гаются пренебрежимо малыми. Использование величин $\tau_1, \tau_2, \dots, \tau_n$ в качестве «весов» вершин графа $\Gamma(M, U)$ позволяет использовать известную задачу определения длин кратчайшего и наиболее длительного пути в ориентированном графе.

Как и в классическом методе Данцига для решения этой задачи преобразуем граф $\Gamma(M, U)$ в сеть, введя две дополнительные вершины H и K .

Начальная вершина H соединена дугами с вершинами графа Γ , имеющими полустепени захода, равные нулю, т.е. не имеющие входных дуг. Конечная вершина K связана дугами с вершинами графа, имеющими полустепени исхода, равные нулю, т.е. не имеющие выходных дуг.

Для дополнительных вершин сети предполагаются следующие значения их «весов»:

$$\tau_H = 0; \quad \tau_K = 0.$$

Для определения значений t_{\min} и t_{\max} преобразуем формулу Данцига:

$$D_j = \min_i \{D_i + d_{ij}\}, \quad D_H = 0,$$

где d_{ij} — «длина» дуги (i, j) графа, в рекуррентное соотношение, которое в общем виде записывается как:

$$t_j = \underset{i \in (0, j-1)}{\text{extr}} \{t_i + \tau_j\}, \quad j = \overline{(0, n+1)}.$$

В этой формуле предполагается, что:

$$\tau_H = \tau_0; \quad \tau_K = \tau_{n+1}.$$

Из предлагаемого соотношения следует, что время выполнения модулей до вершины m_j графа можно вычислить только после того, как будут определены времена t_i выполнения всех предшествующих модулей, связанных дугами с этой вершиной. Тогда искомые параметры t_{\min} и t_{\max} времени работы КП получаются как:

$$t_{\text{extr}} = t_{n+1}.$$

Пример.

Пусть рассматриваемый КП состоит из множества прикладных модулей $M = \{m_1, m_2, \dots, m_5\}$.

Сеть их связи при выполнении функционала КП, построенная по приведенному выше правилу на основе заданного графа $\Gamma(M, U)$, представлена на Рис. 67.

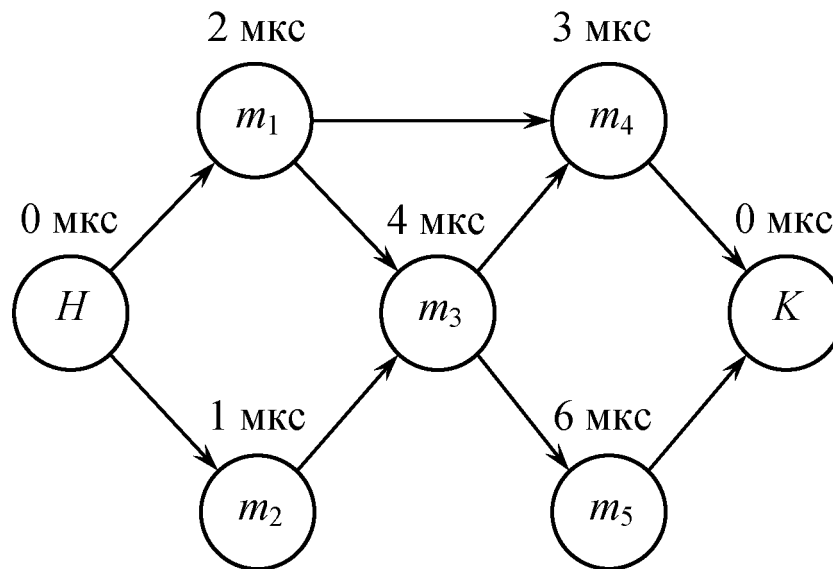


Рис. 67

На этом рисунке приведены в качестве «весов» вершин сети следующие гипотетические значения времени:

$$\tau_H = \tau_0 = 0 \text{ мкс}; \quad \tau_1 = 2 \text{ мкс}; \quad \tau_2 = 1 \text{ мкс};$$

$$\tau_3 = 4 \text{ мкс}; \quad \tau_4 = 3 \text{ мкс}; \quad \tau_5 = 6 \text{ мкс}; \quad \tau_K = \tau_6 = 0 \text{ мкс}.$$

Вычисление величин t_{\min} включает в себя следующие этапы:

$$t_1 = \min\{t_0 + \tau_1\} = 2 \text{ мкс};$$

$$t_2 = \min\{t_0 + \tau_2\} = 1 \text{ мкс};$$

$$t_3 = \min\{t_1 + \tau_3; t_2 + \tau_3\} = \min\{2 + 4; 1 + 4\} = 5 \text{ мкс};$$

$$t_4 = \min\{t_1 + \tau_4; t_3 + \tau_4\} = \min\{2 + 3; 5 + 3\} = 5 \text{ мкс};$$

$$t_5 = \min\{t_3 + \tau_5; t_2 + \tau_3\} = \min\{5 + 6\} = 11 \text{ мкс};$$

$$t_6 = \min\{t_4 + \tau_6; t_5 + \tau_6\} = \min\{5; 11\} = 5 \text{ мкс};$$

Таким образом минимальное значение времени работы КП будет равно:

$$t_{\min} = t_6 = t_K = 5 \text{ мкс.}$$

Для определения величины t_{\max} имеем следующую последовательность шагов:

$$t_1 = \max\{t_0 + \tau_1\} = 2 \text{ мкс;}$$

$$t_2 = \max\{t_0 + \tau_2\} = 1 \text{ мкс;}$$

$$t_3 = \max\{t_1 + \tau_3; t_2 + \tau_3\} = \max\{2 + 4; 1 + 4\} = 6 \text{ мкс;}$$

$$t_4 = \max\{t_1 + \tau_4; t_3 + \tau_4\} = \max\{2 + 3; 6 + 3\} = 9 \text{ мкс;}$$

$$t_5 = \max\{t_3 + \tau_5; t_2 + \tau_3\} = \max\{5 + 6\} = 12 \text{ мкс;}$$

$$t_6 = \max\{t_4 + \tau_6; t_5 + \tau_6\} = \max\{5; 12\} = 12 \text{ мкс;}$$

Отсюда имеем, что:

$$t_{\max} = t_6 = t_K = 12 \text{ мкс.}$$

Приведенную выше модель КП можно применить для планирования комплексного (общего) тестирования, в процессе которого должны быть проверены во взаимодействии все прикладные модули множества $M = \{m_1, m_2, \dots, m_n\}$.

Использование для решения этой задачи полного множества тестов T_r , $r = (\overline{1, k})$ потребует значительных затрат времени на их формирование и в итоге на тестирование КП.

Для сокращения таких затрат будем применять формализм «задачи о минимальном покрытии» [Кристофидес Н.], решение которой позволяет провести проверку всех модулей КП минимальным числом тестов.

Введем булевские переменные:

$$z_r = \begin{cases} 1, & \text{если выбран маршрут тестирования } l_r \in L; \\ 0, & \text{в противном случае.} \end{cases}$$

Эти переменные должны удовлетворять условию:

$$z_r \in \{0; 1\}, \quad r = (\overline{1, k}).$$

Будем считать, что отношение $R_2^{-1} \subseteq M \times L$ описывается матрицей $A = [a_{ir}]_{n \times k}$, отражающей факт включения модулей

$m_i \in M$ в маршруты $l_r \in L$. Тогда условия того, что каждый модуль КП будет принадлежать хотя бы одному тестирующему маршруту, будут иметь вид:

$$\sum_{r=1}^k a_{ir} z_r \geq 1, \quad i = (\overline{1, n}).$$

Общее число получаемых тестирующих маршрутов (ТМ) с учетом булевского характера переменных z_r , $r = (\overline{1, k})$ определяется выражением:

$$N_{\text{ТМ}} = \sum_{r=1}^k z_r \rightarrow \min_{z_r}.$$

Сформированная простейшая модель оптимизации выбора тестов для КП может быть реализована с использованием специализированных алгоритмов для заданной матрицы $[a_{ir}]$ [Кристофидес Н.] или численных методов решения задач целочисленного (дискретного) программирования [Сигал И.Х., Иванова А.В.].

Отметим, что эта задача оптимизации может иметь неединственное решение [Кристофидес Н.]. Для получения единственного решения заменим в ней целевую функцию $N_{\text{ТМ}}$ на критерий оптимальности вида:

$$\Theta = \sum_{r=1}^k \mu_r z_r \rightarrow \min_{z_r}.$$

Здесь параметр $\mu_r = |X_r|$ обозначает мощность подмножества $X_r \subseteq X$ исходных данных (количество параметров и их числовые значения), необходимых для тестирования КП по маршруту $l_r \in L$. Этот критерий косвенным образом описывает общие затраты времени на подготовку тестовых данных.

Пример.

Для графа $\Gamma(M, U)$, приведенного на Рис. 67, полное множество тестов для КП из предыдущего Примера записывается как:

$$\begin{aligned}T_1 &= \{X_1, m_1, m_4, Y_1\}; & T_2 &= \{X_2, m_1, m_3, m_4, Y_3\}; \\T_3 &= \{X_3, m_1, m_3, m_5, Y_3\}; & T_4 &= \{X_4, m_2, m_3, m_4, Y_4\}; \\T_5 &= \{X_5, m_2, m_3, m_5, Y_5\}.\end{aligned}$$

Матрица A , описывающая отношение R_2^{-1} будет иметь вид:

$$A = \begin{matrix} & l_1 & l_2 & l_3 & l_4 & l_5 \\ \begin{matrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \end{matrix} & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{matrix}.$$

Применение к этой матрице одного из алгоритмов, описанных в работе [Кристофидес Н.], дает два варианта минимального покрытия вершин графа (M, U) :

- 1) $L_2 = \{m_1, m_3, m_4\}$, $L_5 = \{m_2, m_3, m_5\}$;
- 2) $L_3 = \{m_1, m_3, m_5\}$, $L_4 = \{m_2, m_3, m_4\}$.

Таким образом, получено два оптимальных решения задачи, которые с использованием введенных выше булевских переменных представляются как:

- 1) $z_1^{(1)} = 0$; $z_2^{(1)} = 1$; $z_3^{(1)} = z_4^{(1)} = 0$; $z_5^{(1)} = 1$;
- 2) $z_1^{(2)} = z_2^{(2)} = 0$; $z_3^{(2)} = 1$; $z_4^{(2)} = 1$; $z_5^{(2)} = 0$.

Из приведенных вариантов Разработчик КП может выбрать наиболее приемлемый с практической точки зрения вариант его тестирования.

Пусть для рассматриваемого КП оценки мощности множеств X_r , $r = \overline{1, 5}$ в количестве входящих в них параметров (пар.) имеют следующие значения:

$$\begin{aligned}\mu_1 &= 15 \text{ пар.}; \quad \mu_2 = 30 \text{ пар.}; \quad \mu_3 = 20 \text{ пар.}; \\ \mu_4 &= 30 \text{ пар.}; \quad \mu_5 = 40 \text{ пар.}\end{aligned}$$

Тогда значения целевой функции $\Theta(z_1, z_2, \dots, z_5)$ для полученных вариантов тестов (T_2, T_5) и (T_3, T_4) будут соответственно равны:

$$\Theta_1 = \mu_2 + \mu_5 = 70 \text{ пар.}; \quad \Theta_2 = \mu_3 + \mu_4 = 50 \text{ пар.}$$

Отсюда следует, что с точки зрения минимизации затрат на подготовку тестов оптимальным является применение тестов T_3 и T_4 .

Отметим, что при использовании полного множества T_1, T_2, \dots, T_5 тестов общее количество подготавливаемых данных будет равно:

$$\Theta_{\text{пол}} = \sum_{r=1}^5 \mu_r = 135 \text{ пар.}$$

Динамическую модель функционирования КП будем строить с использованием формализма марковских случайных процессов с дискретным множеством состояний и с непрерывным временем их смены [Вентцель Е.С.].

Рассмотрение процесса работы КП как случайного процесса объясняется тем, что в процессе его эксплуатации элементы множества X входных данных (наименований параметров) имеют различные числовые значения, которые вызывают различные, в общем случае случайные, затраты времени на выполнение прикладных модулей КП. Случайным можно считать и состав элементов множества X , вводимых в КП при каждом его применении.

Кроме этого случайными будем считать моменты времени применения КП при реализации соответствующих ПИТ.

Для структуры КП, представленного на Рис. 58, введем следующие состояния процесса его функционирования в любой момент времени:

E_0 – работа управляющего модуля;

.....

E_i – работа i -го прикладного модуля;

.....

E_n – работа n -го прикладного модуля.

Будем считать, что смена этих состояний осуществляется с интенсивностями λ_{ki} , где $k, i \in (\overline{0, n})$.

Граф связи состояний КП приведен на Рис. 68.

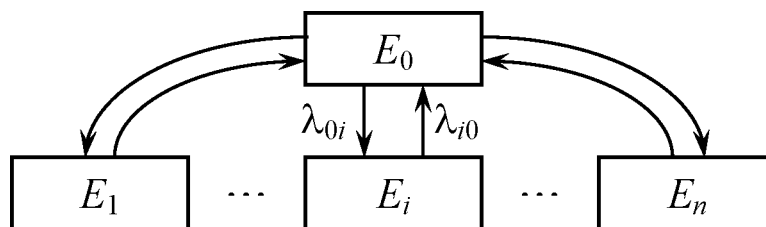


Рис. 68

Введем в рассмотрение вероятности $p_j(t)$ того, что в момент времени t процесс функционирования КП находится в состоянии E_j , $j = (\overline{0, n})$.

Динамику этих вероятностей будем описывать с использованием уравнений Колмогорова. Используя правила их построения [Вентцель Е.С., стр. 199], получаем следующую систему линейных дифференциальных уравнений:

$$\begin{aligned}
 \dot{p}_0 &= -p_0 \sum_{i=1}^n \lambda_{0i} + \lambda_{10}p_1 + \lambda_{20}p_2 + \dots + \lambda_{n0}p_n; \\
 \dot{p}_1 &= -\lambda_{10}p_1 + \lambda_{01}p_0; \\
 &\dots\dots\dots \\
 \dot{p}_i &= -\lambda_{i0}p_i + \lambda_{0i}p_0; \\
 &\dots\dots\dots \\
 \dot{p}_n &= -\lambda_{n0}p_n + \lambda_{0n}p_0, \quad t \in [t_0, t_k].
 \end{aligned} \tag{109}$$

где $\dot{p}_j = \frac{dp_j}{dt}$ – производные по времени, $j = (\overline{0, n})$.

Начальные условия для этой системы записываются как:

$$p_0(t_0) = 1; \quad p_i(t_0) = 0, \quad i = (\overline{1, n}). \quad (110)$$

Эти выражения означают, что в момент времени $t = t_0$ начала работы КП запускается его управляющий модуль, который в последующем с интенсивностями λ_{0i} вызывает на выполнение прикладные модули КП. После работы каждого из таких модулей с интенсивностями λ_{i0} процесс функционирования КП возвращается в состояние E_0 .

Условие того, что этот процесс в любой момент времени должен находиться в одном из состояний E_0, E_1, \dots, E_n имеют вид:

$$\sum_{j=0}^n p_j(t) = 1, \quad t \in [t_0, t_k]. \quad (111)$$

Применим следующий прием для учета этого условия при решении задачи Коши (109), (110).

Выделим из равенства (111) функцию:

$$p_0(t) = 1 - \sum_{j=1}^n p_j(t), \quad (112)$$

отбросим первое уравнение системы (110) и заменим в оставшихся уравнениях вероятность p_0 на правую часть соотношения (112).

Тогда решаемая задача Коши примет вид:

$$\begin{aligned} \dot{p}_i &= -\lambda_{i0} p_i + \lambda_{0i} \left(1 - \sum_{j=1}^n p_j \right), \\ p_i(t_0) &= 0, \quad i = (\overline{1, n}). \end{aligned}$$

Согласно работе [Вентцель Е.С.], для пуассоновских потоков смены состояний марковского процесса входящие в эту

задачу интенсивности переходов могут быть представлены формулами вида:

$$\lambda_{0i} = \frac{1}{T_0}; \quad \lambda_{i0} = \frac{1}{T_i}, \quad i = (\overline{1, n}), \quad (113)$$

где T_0 – средние затраты времени при работе управляющего модуля КП; T_i – среднее время выполнения i -го прикладного модуля КП, $i = (\overline{1, n})$.

В этом случае приведенная выше задача запишется как:

$$\begin{aligned} \dot{p}_i &= -\frac{p_i}{T_i} + \frac{1}{T_0} \left(1 - \sum_{j=1}^n p_j \right), \quad t \in [t_0, t_k]; \\ p_i(t_0) &= 0, \quad i = (\overline{1, n}). \end{aligned} \quad (114)$$

Таким образом, динамическая вероятностная модель функционирования КП описывается выражениями (112) и (114). С использованием этой модели можно изучить ряд свойств рассматриваемого КП. Например, определить из следующих выражений наивероятнейшие моменты времени t_i , в которые выполняются i -е прикладные модули КП:

$$t_i = \arg \max_{t \in [t_0, t_k]} p_i(t), \quad i = (\overline{1, n}). \quad (115)$$

Можно с использованием условия:

$$\max_{i=(\overline{0, n})} |p_i(t + \Delta t) - p_i(t)| \leq \varepsilon \quad (116)$$

определить момент времени $t_{уст} \in [t_0, t_k]$ для достаточно малого ε выхода КП на установившийся режим его работы.

Пример.

Пусть моделируемый КП на интервале времени $[t_0, t_k] = [0; 60]$ с. состоит из $n = 3$ прикладных модулей. Будем считать, что путем обработки соответствующих статистических данных для всех задач, решаемых с помощью этого КП, получены следующие значения:

$$T_0 = 1 \text{ с.}; \quad T_1 = 5 \text{ с.}; \quad T_2 = 15 \text{ с.}; \quad T_3 = 10 \text{ с.}$$

Тогда задача Коши (114) примет вид:

$$\dot{p}_1 = -\frac{p_1}{5} + (1 - p_1 - p_2 - p_3);$$

$$\dot{p}_2 = -\frac{p_2}{15} + (1 - p_1 - p_2 - p_3);$$

$$\dot{p}_3 = -\frac{p_3}{10} + (1 - p_1 - p_2 - p_3);$$

$$p_1(0) = 0; \quad p_2(0) = 0; \quad p_3(0) = 0.$$

Численное решение этой задачи и применение выражений (115) и (116) дало следующие результаты:

1. Наивероятнейшие моменты времени выполнения прикладных модулей КП имеют вид:

$$t_1 = 1,8 \text{ с}; \quad p_{1 \max} = 0,279;$$

$$t_2 = 60 \text{ с}; \quad p_{2 \max} = 0,483;$$

$$t_3 = 15 \text{ с}; \quad p_{3 \max} = 0,344.$$

2. Моменты выхода КП на установившийся режим работы при $\varepsilon = 0,001$:

$$t_{\text{уст}} = \max\{t_{\text{уст}}^{(0)}, t_{\text{уст}}^{(1)}, t_{\text{уст}}^{(2)}, t_{\text{уст}}^{(3)}\} = \max\{18, 36, 54, 48\} = 54 \text{ с.}$$

3. Установившиеся значения вероятностей:

$$p_0^* = 0,033; \quad p_1^* = 0,161; \quad p_2^* = 0,483; \quad p_3^* = 0,323.$$

Отсюда можно сделать практический вывод, что работа управляющего модуля КП занимает 3,3% времени эксплуатации комплекса, 16,1% времени занимает выполнение модуля 1, 48,3% – модуля 2 и 32,3% – модуля 3.

Из приведенного примера следует наличие у процесса функционирования КП установившегося режима его работы.

Статическая вероятностная модель работы КП в таком режиме для определения установившихся значений вероятно-

стей p_j^* , $j = (\overline{0, n})$ может быть получена при значениях $\dot{p}_j \equiv 0$, $j = (\overline{0, n})$ из уравнений (109).

Отбросим первое уравнение получаемой системы линейных алгебраических уравнений и проведем в них замену вида (112). Тогда формируемая модель с использованием условия нормировки вероятностей p_j^* примет следующий вид:

$$-\frac{p_i^*}{T_i} + \frac{p_0^*}{T_0} = 0, \quad i = (\overline{1, n});$$

$$\sum_{j=0}^n p_j^* = 1.$$

Из первых n уравнений этой системы получим выражения:

$$p_i^* = \frac{T_i}{T_0} p_0^*, \quad i = (\overline{1, n}). \quad (117)$$

Подставляя их в условия нормировки вероятностей p_j^* , $j = (\overline{0, n})$, определим значение p_0^* из полученного уравнения:

$$p_0^* + \frac{p_0^*}{T_0} \sum_{j=1}^n T_j = 1,$$

следующим образом:

$$p_0^* = \frac{T_0}{\sum_{j=1}^n T_j}.$$

Подставляя это соотношение в выражения (117), получаем расчетные формулы:

$$p_i^* = \frac{T_i}{\sum_{j=1}^n T_j}, \quad i = (\overline{0, n})$$

для определения установившихся значений вероятностей выполнения всех модулей КП.

Таким образом, сформированная статическая вероятностная модель КП по классификации Лекции №4 относится к аналитически-разрешимым математическим моделям.

В заключение лекции скажем несколько слов о методологии разработки прикладного (функционального) ПО ИС, к которому относятся всевозможные комплексы программ.

В СССР для создания ПО решения сложных научно-технических задач и автоматизированных (информационных) систем различного назначения использовалась группа обязательных для исполнения государственных стандартов (ГОСТ) «Единая система программной документации (ЕСПД)».

Стандарты этой группы регламентировали стадии и этапы создания ПО и их содержание от разработки технического задания (ТЗ) до выпуска программной (ПД) и эксплуатационной (ЭД) документации на готовый программный продукт (изделие).

Основным назначением комплекта документации, которое наряду с описаниями алгоритмов и программ (модулей) включала в себя такие документы как «Руководство системному программисту» и «Руководство программисту», являлось предоставление Заказчику ПО возможностей самостоятельно вносить необходимые изменения при модификации разработанного программного продукта.

В настоящее время практически все зарубежные и отечественные программные продукты передаются Заказчикам (Покупателям) без сколь-нибудь подробного описания их содержания, а их сопровождение и модификация осуществляется Разработчиком (Продавцом) за отдельную плату.

Обязательное применение ГОСТ ЕСПД при разработках такой важной компоненты ИС как ПО не допускало современной «анархии» и рыночного произвола Разработчиков (Продавцов) программных продуктов.

Дополнительную «анархию» вносят всевозможные в основном зарубежные псевдонаучные теории разработки ПО, которые предлагают «каскадные», «итерационные» и другие модели процессов разработки ПО, тогда как в ЕСПД были четко обозначены обязательные для всех стадии создания программного продукта от ТЗ до ЭД.

Раздел V. ЗАДАЧИ ОРГАНИЗАЦИИ РАЗРАБОТКИ И ЭКСПЛУАТАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ

В данном разделе представлены математические модели некоторых задач, возникающих при организации эффективных процессов разработки и эксплуатации компонент информационных систем. Эти модели можно рассматривать как первый этап исследований и разработок этой важнейшей проблемы современной прикладной информатики.

19. Модели некоторых задач организации разработки информационных систем

В прикладной информатике математическое моделирование должно применяться не только к формальному описанию элементов информационных систем (ИС) и технологий (ИТ), но и к процессам их разработки. Последнее, на наш взгляд, позволит резко повысить эффективность этих процессов и в конечном итоге и эффективность функционирования ИС.

Практика создания таких систем и технологий выявила значительное число задач, возникающих в процессах их разработки.

Отметим, что формализация постановок и решения этих задач со времен отказа от отечественных подходов к созданию разнообразных автоматизированных систем практически не развивается.

Рассмотрим модели некоторых задач организации разработок ИС и ИТ.

Задача формирования состава распределенного банка данных (РБнД)

Пусть в некоторой корпоративной информационной системе (КИС), включающей в себя множество $N = \{1, 2, \dots, n\}$

подразделений, сформирована совокупность $B = \{B_1, B_2, \dots, B_m\}$ требуемых для их работы баз данных (БД), где $m < n$.

Потребности подразделений КИС в информации, хранящейся в этих БД, будем представлять стохастическим отношением вида:

$$R_1^c \subseteq N \times B.$$

Согласно Лекции №11 матрица A этого отношения состоит из элементов p_{ik} , представляющих собой вероятности того, что при выполнении прикладных информационных технологий (ПИТ) i -го подразделения КИС, будет использована информация из базы данных B_k , $i = (\overline{1, n})$, $k = (\overline{1, m})$. Эти вероятности должны удовлетворять условиям нормировки вида:

$$\sum_{k=1}^m p_{ik} = 1, \quad i = (\overline{1, n}).$$

На практике вероятности p_{ik} могут быть приближенно определены как относительные частоты:

$$p_{ik} \cong \frac{n_{ik}}{\sum_{k=1}^m n_{ik}},$$

где n_{ik} — количество выполнений всех ПИТ в i -м подразделении КИС, использующих информацию, хранящуюся в базе данных B_k .

Требуется найти распределение БД по локальным банкам данных (ЛБНД) подразделений КИС, рассмотренных в Лекции №17.

Для решения этой задачи будем использовать переменные x_{ik} , которые могут принимать следующие значения:

$$x_{ik} = \begin{cases} 1, & \text{если база } B_k \text{ включена в состав } i\text{-го ЛБНД;} \\ 0, & \text{в противном случае.} \end{cases}$$

Отсюда возникают условия вида:

$$x_{ik} \in \{0; 1\}, \quad i = (\overline{1, n}), \quad k = (\overline{1, m}). \quad (115)$$

С учетом этого система равенств:

$$\sum_{k=1}^m x_{ik} = 1, \quad i = (\overline{1, n}), \quad (116)$$

описывает требование, чтобы каждая база B_k была размещена в одном из ЛБНД.

Пусть W_i – объем памяти сервера i -го подразделения, отведенного под ЛБНД (Мбайт); w_k – объем памяти, требуемой базе B_k (Мбайт).

Тогда имеем ограничения вида:

$$\sum_{k=1}^m w_k x_{ik} \leq W_i, \quad i = (\overline{1, n}). \quad (117)$$

Обозначим через $\lambda_i^{(BX)}$ интенсивность обращений за требуемой информацией при выполнении всех ПИТ i -го подразделения КИС (c^{-1}); v_i – максимальный объем одного обращения в i -м подразделении, включающего в себя собственно запрос и данные ответа на него (Мбайт).

Ограничение по пропускной способности системы связи между подразделениями КИС записывается как:

$$\sum_{k=1}^m \sum_{i=1}^n \lambda_i^{(BX)} v_i (1 - x_{ik}) \leq S, \quad (118)$$

где S – скорость передачи данных сети связи КИС (Мбайт/с).

Правая часть этого неравенства означает с учетом условий (115) суммарную скорость потока неудовлетворенных собственными ЛБНД подразделений обращений к информации, хранящейся в других ЛБНД.

Потребуем при формировании РБНД обеспечения для каждого подразделения КИС максимальной автономности рабо-

ты с требуемыми данными, т.е. наличия максимально «полного наполнения» их ЛБнД.

С использованием отношения R^c вероятность того, что i -ое подразделение будет использовать информацию из БД, включенных в его ЛБнД, определяется по формуле:

$$P_i = \prod_{k=1}^m p_{ik}^{x_{ik}}, \quad i = (\overline{1, n}).$$

Входящие в это выражение переменные должны удовлетворять условиям (115) и (116). Вследствие этого данная формула примет следующий вид [Раскин А.Г.]:

$$P_i = \sum_{k=1}^m p_{ik} x_{ik}, \quad i = (\overline{1, n}).$$

Тогда критерий оптимальности формирования РБнД КИС представляется как:

$$Q = \sum_{i=1}^n P_i = \sum_{i=1}^n \sum_{k=1}^m p_{ik} x_{ik} \rightarrow \max_{x_{ik}}. \quad (119)$$

Выражения (115)-(119) представляют модель задачи линейного дискретного (булевского) программирования, которая может быть решена с использованием известных численных методов [Сигал И.Х., Иванова А.В. и др.] или с помощью эвристических алгоритмов [Цегелик Г.Г.].

Отметим, что в последней работе рассматриваются различные постановки задачи оптимального размещения файлов по узлам вычислительной сети.

Результаты решения этой задачи позволяют определить матрицу передач $T = [\theta_{ij}]_{n \times n}$, описывающую взаимодействие в составе РБнД КИС сформированных ЛБнД ее подразделений (см. Лекцию №17).

Пусть x_{ik}^0 – результаты решения задачи (115)-(119). Сведем эти значения в булевскую матрицу $X_0 = [x_{ik}^0]_{n \times n}$, которая может рассматриваться как представление классического (детерминированного) бинарного отношения:

$$R_2 \subseteq N \times B,$$

описывающего распределение БД по подразделениям КИС.

Используя элементы матриц A и X_0 , определим искомые вероятности передачи обращений за потребными данными, инициированными в i -м подразделении к j -му подразделению КИС, по формуле [Цигелик Г.Г.]:

$$\theta_{ij} = \sum_{r=1}^m p_{ir} x_{rj}^0, \quad i, j = (\overline{1, n}). \quad (120)$$

При этом требуемые в модели РБНД условия нормировки этих вероятностей:

$$\sum_{j=1}^n \theta_{ij} = 1, \quad i = (\overline{1, n})$$

выполняются.

Формулу (120) в матричном виде можно представить соотношением:

$$T = AX_0^*,$$

где X_0^* – транспонированная матрица X_0 результатов решения задачи (115)-(119).

Это соотношение определяет стохастическое бинарное отношение:

$$R_3^c \subseteq N \times N,$$

которое описывает взаимодействие ПИТ подразделений рассматриваемой КИС.

Пример.

Воспользуемся исходными данными из монографии [Цигелик Г.Г.].

Пусть КИС состоит из $n = 3$ подразделения, для деятельности которых необходимы $m = 7$ баз данных. Будем считать известными значения вероятностей p_{ik} , описывающих отношение R_1^c :

$$A = \begin{matrix} & \begin{matrix} B_1 & B_2 & B_3 & B_4 & B_5 & B_6 & B_7 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0,05 & 0,30 & 0,15 & 0,25 & 0,10 & 0,06 & 0,09 \\ 0,40 & 0,10 & 0,05 & 0,08 & 0,12 & 0,10 & 0,15 \\ 0,15 & 0,07 & 0,40 & 0,03 & 0,10 & 0,15 & 0,10 \end{bmatrix} \end{matrix}.$$

Предположим, что оптимальное решение задачи (115)-(119) имеет вид:

$$\begin{aligned} x_{12}^0 = x_{14}^0 = x_{15}^0 = 1; & \quad x_{11}^0 = x_{13}^0 = x_{16}^0 = x_{17}^0 = 0; \\ x_{21}^0 = x_{26}^0 = x_{27}^0 = 1; & \quad x_{22}^0 = x_{23}^0 = x_{24}^0 = x_{25}^0 = 0; \\ x_{33}^0 = 1; & \quad x_{31}^0 = x_{32}^0 = x_{34}^0 = x_{35}^0 = x_{36}^0 = x_{37}^0 = 0. \end{aligned}$$

Матрица X^* имеет вид:

$$X^* = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{matrix}.$$

Из этой матрицы следует, что в ЛБНД 1 первого подразделения КИС входят базы данных B_2, B_4, B_5 , в ЛБНД 2 – B_1, B_6, B_7 , а в ЛБНД 3 – только B_3 .

Применяя формулу (120), получаем искомую матрицу передач РБнД КИС вида:

$$T = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix} = \begin{bmatrix} 0,65 & 0,20 & 0,15 \\ 0,30 & 0,65 & 0,05 \\ 0,20 & 0,40 & 0,40 \end{bmatrix}.$$

Заметим, что наличие в ней достаточно больших значений вероятностей θ_{ii} , $i = (\overline{1, 3})$ говорит о том, что при рассмотренных исходных данных подразделения КИС в основном используют свои ЛБнД.

Задача выбора средств разработки ПО прикладных задач

Пусть на ИТ-рынке представлено K образцов средств разработки ПО определенного функционального назначения (платформы, системы программирования, кросс-системы и т.п.). Каждое k -е средство имеет определенную стоимость c_k (д.е.) и количество f_k реализуемых функций (задач), $k = (\overline{1, K})$.

Требуется выбрать из каталога средств наиболее эффективное средство, имеющее минимальную цену и максимальное число функций.

Для решения этой задачи введем булевские переменные:

$$x_k \in \{0; 1\}, \quad k = (\overline{1, K}),$$

определяющие выбор конкретного средства.

Условие выбора одного из K средств имеет следующий стандартный вид:

$$\sum_{k=1}^K x_k = 1. \quad (121)$$

Согласно приведенным выше требованиям сформируем целевые функции задачи:

$$C = \sum_{k=1}^K c_k x_k \rightarrow \min_{x_k}; \quad F = \sum_{k=1}^K f_k x_k \rightarrow \max_{x_k}.$$

Представленные выражения описывают модель двухкритериальной задачи оптимального выбора средства разработки ПО.

Как известно, многокритериальные задачи оптимизации не имеют единственного решения [Подиновский В.В., Ногин В.Д.].

Будем определять решение нашей задачи сведением к однокритериальной задаче оптимизации с использованием подхода, принятого в инженерной практике, и принципа оптимальности Парето [Подиновский В.В., Ногин В.Д.].

При первом подходе при решении задачи применяется вспомогательная целевая функция, описывающая нелинейную свертку критериев вида:

$$W(x_1, x_2, \dots, x_K) = \frac{C(x_1, x_2, \dots, x_K)}{F(x_1, x_2, \dots, x_K)} = \frac{\sum_{k=1}^K c_k x_k}{\sum_{k=1}^K f_k x_k} \rightarrow \min_{x_k}.$$

С использованием условия (121) эту функцию можно представить в следующем виде:

$$W = \sum_{k=1}^K \frac{c_k}{f_k} x_k \rightarrow \min_{x_k}.$$

Отсюда следует, что номер k^0 оптимального средства разработки ПО выбирается из рассматриваемого каталога с использованием соотношения:

$$k^0 = \arg \min_{k=(1, K)} \{c_k / f_k\}. \quad (122)$$

При использовании второго подхода дискретное парето-оптимальное множество в пространстве критериев решаемой задачи с номерами k_j , $j = \overline{(1, K)}$ рассматриваемых средств имеет вид, представленный на Рис. 69.

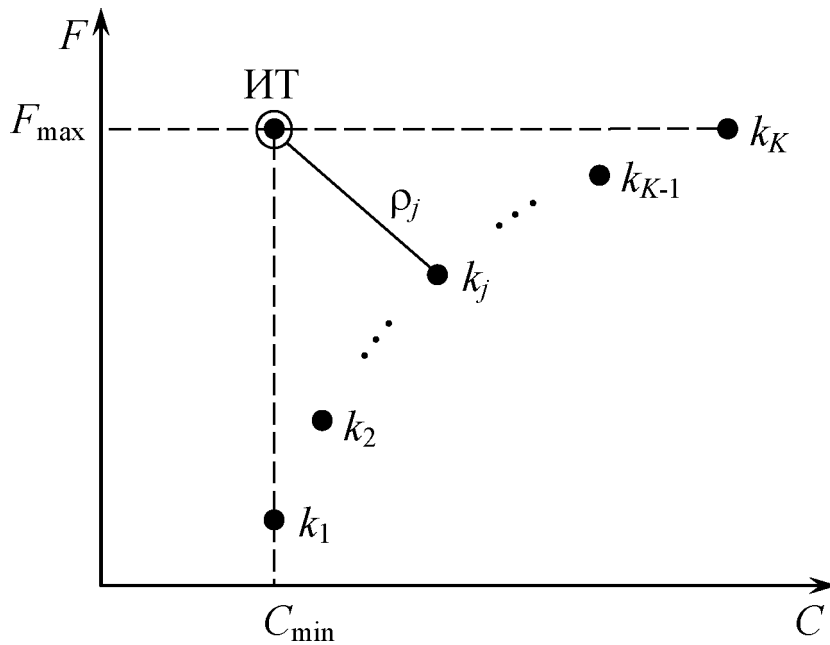


Рис. 69

Это множество в форме совокупности значений k_j , c_{k_j} , f_{k_j} , $j = (\overline{1, K})$ может быть выдана на АРМ ЛПР по приобретению средств разработки ПО для выбора им компромиссного решения на основе его опыта и интуиции.

Единственное оптимальное решение может быть получено с использованием метода «идеальной точки», проиллюстрированном в Лекции №3. В нашем случае координаты такой точки в пространстве критериев задачи будут равны (C_{\min}, F_{\max}) , где $C_{\min} = \min_{k=(\overline{1, K})} c_k$, $F_{\max} = \max_{k=(\overline{1, K})} f_k$.

Тогда номер k^0 оптимального средства разработки ПО определяется из выражения:

$$k^0 = \arg \min_{j=(\overline{1, K})} \rho_j = \arg \min \sqrt{(c_{k_j} - C_{\min})^2 + (f_{k_j} - F_{\max})^2}, \quad (123)$$

где ρ_j – расстояние от ИТ до точек паретооптимального множества решений в пространстве критериев.

Пример.

Пусть требуется выбрать средство разработки ERP-технологий из каталога, содержащего $K = 4$ образца со следующими условными стоимостными (млн. руб.) и функциональными показателями:

$$c_1 = 1,160 \text{ млн. руб.}; \quad c_2 = 1,814 \text{ млн. руб.};$$

$$c_3 = 4,500 \text{ млн. руб.}; \quad c_4 = 3,200 \text{ млн. руб.};$$

$$f_1 = 8; \quad f_2 = 13; \quad f_3 = 20; \quad f_4 = 16.$$

Используя соотношение (122), имеем:

$$\begin{aligned} k^0 &= \arg \min \left\{ \frac{1,160}{8}; \frac{1,814}{13}; \frac{4,500}{20}; \frac{3,200}{16} \right\} = \\ &= \arg \min \{0,145; 0,139; 0,225; 0,200\} = 2. \end{aligned}$$

Отсюда следует, что необходимо использовать второе средство разработки системы как оптимальное по распространенному на практике критерию «цена-качество».

Паретооптимальное множество задачи приведено на Рис. 70.

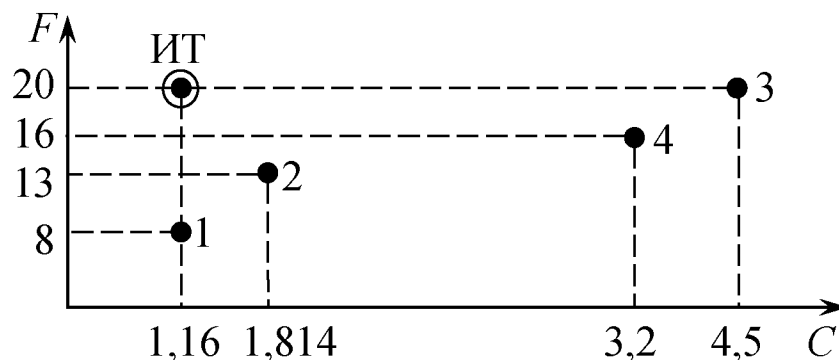


Рис. 70

Координаты «идеальной точки» (ИТ) будут равны:

$$(C_{\min}, F_{\max}) = (1,16; 20).$$

Определим расстояния от этой точки до показателей каждого образца:

$$\rho_1 = \sqrt{(1,16 - 1,16)^2 + (8 - 20)^2} = 12;$$

$$\rho_2 = \sqrt{(1,814 - 1,16)^2 + (13 - 20)^2} = 7,03;$$

$$\rho_3 = \sqrt{(4,50 - 1,16)^2 + (20 - 20)^2} = 3,34;$$

$$\rho_4 = \sqrt{(3,20 - 1,16)^2 + (16 - 20)^2} = 4,49.$$

Тогда согласно (123) получаем:

$$k^0 = \arg \min\{12; 7,03; 3,34; 4,49\} = 3.$$

Таким образом, при использовании второго подхода необходимо выбрать средство 3.

Задача подбора специалистов для выполнения ИТ-проектов

Пусть для разработки ПО определенных технологий некоторая организация приобрела n средств их создания. Для их применения нужны ИТ-специалисты, знакомые с возможностями этих средств и имеющие навыки их эксплуатации.

Будем считать, что имеется m претендентов на работу в этой организации. При этом каждый i -й ИТ-специалист претендует на оплату его труда в размере c_j руб., $j = (\overline{1, m})$, но владеет только определенным подмножеством из n средств разработки ПО. Требуется выбрать из множества претендентов конкретных ИТ-специалистов, которые в совокупности владели бы всеми имеющимися средствами разработки и требовали минимальных затрат на их зарплату.

Введем следующие обозначения:

$S = \{s_1, s_2, \dots, s_n\}$ – множество средств разработки ПО;

$\Pi = \{\pi_1, \pi_2, \dots, \pi_m\}$ – множество ИТ-специалистов, претендующих на работу в данной организации.

Возможности этих специалистов будем описывать бинарным отношением:

$$R \subseteq S \times \Pi.$$

Матричное представление $A = [a_{ij}]_{n \times m}$ этого отношения имеет следующие элементы:

$$a_{ij} = \begin{cases} 1, & \text{если } i\text{-й специалист владеет} \\ & j\text{-м средством разработки ПО;} \\ 0, & \text{в противном случае.} \end{cases}$$

Сформулированную выше задачу будем решать с использованием формализма «задачи о минимальном покрытии» [Кристофидес Н.], которая описывается математической моделью вида:

$$\begin{aligned} C &= \sum_{j=1}^m c_j x_j \rightarrow \min_{x_j}; \\ \sum_{j=1}^m a_{ij} x_j &\geq 1, \quad i = (\overline{1, n}); \\ x_j &\in \{0; 1\}, \quad j = (\overline{1, m}). \end{aligned}$$

Здесь последние ограничения указывают на булевскую природу переменных модели. С их учетом вторые соотношения модели описывают требования, чтобы отобранные ИТ-специалисты владели в совокупности всеми имеющимися средствами разработки.

Приведенная выше задача решается с помощью алгоритмов, приведенных в книге [Кристофидес Н.] или численных методов булевого (дискретного) программирования, описанных в работе [Сигал И.Х., Иванова А.В. и др.]. Кроме этого задача может быть решена специальным эвристическим алгоритмом типа «жадных алгоритмов» [Компьютер и задачи выбора].

Пример.

Пусть некоторая организация приобрела $m = 7$ средств разработки ПО. Заявки на работу в ней подали $n = 5$ специалистов со следующими требованиями на оплату труда:

$$c_1 = 80 \text{ тыс. руб.}; \quad c_2 = 40 \text{ тыс. руб.}; \quad c_3 = 35 \text{ тыс. руб.};$$

$$c_4 = 45 \text{ тыс. руб.}; \quad c_5 = 50 \text{ тыс. руб.}$$

Будем считать, что матрица A , описывающая знания средств разработки ИТ-специалистами, имеет вид:

$$A = \begin{matrix} & \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \\ \begin{matrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \end{matrix} & \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{matrix}$$

С использованием приведенных исходных данных модель задачи подбора ИТ-специалистов конкретизируется как:

$$C = 80x_1 + 40x_2 + 35x_3 + 45x_4 + 50x_5 \rightarrow \min;$$

$$x_1 + x_3 + x_4 \geq 1; \quad x_1 + x_2 \geq 1; \quad x_2 \geq 1;$$

$$x_1 + x_4 \geq 1; \quad x_5 \geq 1; \quad x_2 + x_3 + x_5 \geq 1; \quad x_4 + x_5 \geq 1;$$

$$x_1 \in \{0; 1\}, \quad x_2 \in \{0; 1\}, \dots, x_5 \in \{0; 1\}.$$

Оптимальное решение этой задачи, полученное по матрице A с использованием одного из алгоритмов работы [Кристофидес Н.], записывается как:

$$x_1^0 = x_5^0 = 0; \quad x_2^0 = x_3^0 = x_4^0 = 1; \quad C_{\min} = 120 \text{ тыс. руб.}$$

Таким образом, на работу в организации необходимо взять специалистов π_2 , π_3 и π_4 , каждый из которых владеет тремя различными средствами разработки ПО.

Задача формирования состава прикладных модулей комплекса программ (КП)

Пусть разрабатываемый КП предназначен для реализации F функций (задач, заданий) некоторой прикладной информационной технологии. Будем считать, что были разработаны и автономно отлажены F программ, реализующих требуемые

функции. При этом в процессе тестирования этих программ были измерены следующие их «пространственно-временные» характеристики:

v_f – объем f -й программы;

τ_f – время выполнения каждой программы $f = (\overline{1, F})$.

Обозначим через N искомое число прикладных модулей КП, в которые необходимо скомпоновать F имеющихся программ. Для определения такой компоновки будем использовать булевские переменные, которые принимают следующие значения:

$$x_{fn} = \begin{cases} 1, & \text{если программа } f \text{ включена} \\ & \text{в модуль } n; \\ 0, & \text{в противном случае.} \end{cases}$$

Отсюда вытекают ограничения вида:

$$x_{fn} \in \{0; 1\}, \quad f = (\overline{1, F}), \quad n = (\overline{1, N}). \quad (124)$$

Условия того, что каждая программа должна входить в определенный прикладной модуль, имеют вид:

$$\sum_{f=1}^F x_{fn} = 1, \quad n = (\overline{1, N}). \quad (125)$$

Ограничение на число формируемых модулей КП записывается как:

$$2 \leq N < F, \quad N - \text{целое число.} \quad (126)$$

Ограничения на объем каждого модуля представим следующими неравенствами:

$$\sum_{f=1}^F v_f x_{fn} \geq v^*, \quad n = (\overline{1, N}), \quad (127)$$

в правых частях которых использовано значение:

$$v^* = \max_{f=(\overline{1, F})} \{v_f\}.$$

Будем считать, что дан граф $G(F, U)$, описывающий взаимосвязи функций $f \in (\overline{1, F})$. Тогда требование того, что в каждый модуль необходимо включать связанные между собой программы k и r , формализуется как:

$$\sum_{k=1}^F \sum_{r=1}^F a_{kr} x_{kn} x_{rn} = 1, \quad n = (\overline{1, N}). \quad (128)$$

где a_{kr} — элементы матрицы смежности A графа $G(F, U)$.

Суммарное время выполнения всех модулей КП определяется выражением вида:

$$T_1(x, N) = \sum_{f=1}^F \sum_{n=1}^N \tau_f x_{fn}.$$

Примем, что время модификации КП будет обратно пропорционально числу входящих в него прикладных модулей. Это означает, что чем больше число N , тем меньше времени затрачивает разработчик КП на поиск модифицируемой функции и внесения в нее требуемых изменений.

Оценку затрат времени на модификацию КП представим следующей формулой:

$$T_2(N) = \frac{\tau^*}{N},$$

где $\tau^* = \sum_{f=1}^F \tau_f$ — суммарное быстроедействие всех разработанных программ.

Введем целевую функцию задачи формирования оптимального состава прикладных модулей разрабатываемого КП:

$$\begin{aligned} T(x, N) &= T_1(x, N) + T_2(N) = \\ &= \sum_{f=1}^F \sum_{n=1}^N \tau_f x_{fn} + \frac{\tau^*}{N} \rightarrow \min_{x, N}. \end{aligned} \quad (129)$$

Заметим, что эта функция имеет минимум по целочисленной переменной N , т.к. с ее ростом в полуинтервале $[2, F)$ первое слагаемое возрастает, а второе – убывает.

Таким образом, выражения (124)-(129) представляют собой модель задачи нелинейного целочисленного программирования, в которой одна из переменных входит в целевую функцию в виде верхнего предела суммирования. Кроме этого, искомое значение N делает переменным количество ограничений (124)-(128).

Для учета этих особенностей задачи, которые практически не отражены в существующей литературе, предлагается осуществить ее решение итерационным процессом путем последовательного задания значений N , удовлетворяющих условию (126), и применения существующих численных методов решения задач нелинейного булевого программирования. Такой процесс продолжается до получения минимального значения функции $T(x, N)$.

Для сокращения перебора можно для нахождения очередного значения N использовать известные методы одномерной оптимизации, такие как метод дихотомии, метод Фибоначи и др.

В заключение лекции отметим, что для успешного развития прикладных ИТ необходима разработка теоретических основ их создания с последующей реализацией полученных моделей, методов и алгоритмов в перспективных САПР ПИТ.

20. Некоторые модели процессов эксплуатации информационных систем

Как и в предыдущей лекции можно констатировать факт практического отсутствия в настоящее время работ, формализующих процессы эксплуатации разнообразных ИС и ИТ. Исключением является монография [Зиновьев П.А. и др.], в ко-

торой приводятся математические модели и статистические данные по функционированию разнообразных КИС.

Рассмотрим некоторые задачи, возникающие в практике эксплуатации прикладных информационных технологий и систем.

Задача резервирования программных средств и данных

Резервирование (дублирование) программных средств и данных, используемых при реализации ИТ, является одним из системных методов повышения эффективности эксплуатации ИС.

При эксплуатации ИС присутствует возможность их разрушений, которые приводят к появлению ошибочных результатов применения ИТ, невозможности решения некоторых задач и даже к полному отказу ИС. Восстановление потерянной информации и программ может потребовать значительных затрат времени и средств. Например, для восстановления «больших» баз данных (БД) потребуются трудозатраты порядка нескольких человеко-лет.

Под разрушением программы или файла понимается такое их состояние, при котором в результате потери части или всей содержащейся в них информации КП или БД не могут использоваться ИС для получения требуемых результатов.

Основными причинами их разрушения являются:

- неисправности носителя информации;
- ошибки пользователей;
- ошибки программ;
- результаты информационных атак на ИС.

Для обеспечения сохранности программных модулей и файлов при функционировании автоматизированных систем

применяется их резервирование [Мамиконов А.Г., Кульба В.В., Шелков А.Б.]. В этой фундаментальной работе приводятся факторы, влияющие на процессы разрушения модулей и файлов, которые должны быть «осовременены» с учетом современного состояния программных и аппаратных средств прикладной информатики. В ней же рассматриваются следующие виды резервирования:

- *оперативное резервирование*, при котором резервные копии модулей и файлов используются только для решения ИС требуемых задач;
- *восстановительное резервирование*, при котором часть резервных копий применяется только для восстановления разрушенных рабочих копий, предназначенных для решения текущих задач;
- *резервирование для длительного хранения* программных модулей и файлов.

Для их реализации предлагаются следующие стратегии резервирования:

1. Применение нескольких последовательно используемых копий каждого файла.
2. Использование в качестве копий файла текущих данных его предыстории (предыдущие файлы и файлы изменений).
3. Смешанная стратегия резервирования, состоящая из элементов стратегий 1 и 2.

Основными количественными характеристиками стратегий резервирования являются среднее время до разрушения файла и его копий (предыстории) и вероятность разрушения файла за единичный интервал времени его использования.

В этой работе приводится следующая формула для определения количества ΔN резервных копий при использовании стратегии 1:

$$\Delta N = \left[\left(\frac{\ln(1 - P_{\text{тр}})}{\ln q} \right) \right],$$

где $P_{\text{тр}}$ – требуемое значение вероятности успешного решения задачи с использованием основного файла и его резервных копий; q – вероятность разрушения файла и его копий за единственный интервал времени; $[(\cdot)]$ – операция округления числа (\cdot) до ближайшего целого значения в большую сторону.

Модели оптимального размещения копий файлов распределенной базы данных в вычислительной сети предлагаются в монографии [Цегелик Г.Г.].

Рассмотрим основанный на теории надежности восстанавливаемых систем [Вентцель Е.С.] метод определения числа ΔN резервных копий определенной базы данных (БД), которая используется при реализации некоторой прикладной информационной технологии (ПИТ).

Пусть эта технология выполняется в интервале времени длительностью T час.

Будем считать известной величину λ интенсивности выхода из строя (разрушения) применяемой БД и ее резервных копий, имеющую размерность час⁻¹.

Выход из строя БД до завершения ПИТ может отрицательным образом повлиять на достижение целей ее реализации в ИС. Отсюда возникает задача определения дополнительного числа ΔN резервных копий БД, позволяющего совместно с основной БД технологии достичь этих целей с не менее чем заданной требуемой вероятностью $P_{\text{тр}}$.

Будем считать, что в процессе выполнения ПИТ каждая вышедшая из строя версия БД «мгновенно» заменяется одной из оставшихся резервных копий. Это означает, что затраты времени на замену разрушенной БД ее копией являются пренебрежимо малыми. Тогда успешность выполнения ПИТ будет зависеть от достаточности числа ΔN на интервале времени $(0, T)$.

Если считать, что случайный поток выхода из строя БД является стационарным пуассоновским потоком с интенсивностью λ , то, следуя работе [Вентцель Е.С.], вероятность достаточности числа $(N + \Delta N)$ экземпляров БД для выполнения запланированной технологии может быть определена как:

$$Q_{\text{дост}}(\Delta N) = \sum_{n=0}^{N+\Delta N} \frac{(\lambda T)^n}{n!} e^{-\lambda T}. \quad (130)$$

Потребуем, чтобы искомое число ΔN такое, что:

$$\Delta N \in \{0, 1, 2, 3, \dots\},$$

удовлетворяло условию вида:

$$Q_{\text{дост}}(\Delta N) \geq P_{\text{тр}}. \quad (131)$$

Представим выражение (130) в следующей форме:

$$Q_{\text{дост}}(\Delta N) = e^{-\lambda T} \sum_{n=0}^N \frac{(\lambda T)^n}{n!} + e^{-\lambda T} \sum_{k=0}^{\Delta N} \frac{(\lambda T)^k}{k!}.$$

Подставляя это выражение в неравенство (131) и проводя несложные преобразования, получаем следующее соотношение для определения величины ΔN :

$$\sum_{k=0}^{\Delta N} \frac{(\lambda T)^k}{k!} \geq P_{\text{тр}} e^{\lambda T} - \sum_{n=0}^N \frac{(\lambda T)^n}{n!}. \quad (132)$$

Сложность решения этого неравенства состоит в том, что искомая величина ΔN является верхним пределом суммирования. Такие задачи обычно решаются последовательным та-

булированием значений левой части неравенства (132) до его выполнения.

Для упрощения проводимых расчетов введем в рассмотрение функцию целочисленного аргумента:

$$\varphi(\Delta N) = \sum_{k=0}^{\Delta N} \frac{(\lambda T)^k}{k!} \quad (133)$$

и параметр:

$$a = P_{\text{тр}} e^{\lambda T} - \sum_{n=0}^1 \frac{(\lambda T)^n}{n!}. \quad (134)$$

Тогда неравенство (132) примет следующий вид:

$$\varphi(\Delta N) \geq a. \quad (135)$$

Вычислительная процедура решения этого неравенства включает в себя предварительное определение значения параметра a по формуле (134) и последовательный расчет значений функции (133) при $\Delta N = 0, \Delta N = 1, \Delta N = 2, \dots$ до выполнения условия (135).

Значение параметра λ , используемого в этой процедуре, может быть получено известными методами теории надежности путем обработки статистики по выходу из строя БД за определенный период времени.

Приближенную оценку значения этого параметра предлагается вычислять как:

$$\lambda = \frac{M}{\tau},$$

где M – число вышедших из строя БД в предшествующий интервал времени эксплуатации ПИТ длительностью τ час.

Использованное предположение о том, что случайное число вышедших из строя БД подчиняется закону Пуассона, позволяет определить по формуле:

$$m_{\Delta N} = \lambda T$$

их среднее (ожидаемое) количество в течение интервала времени длительностью T час.

Пример.

Пусть планируется работа с некоторой БД длительностью $T = 8$ час. При предыдущей работе с БД длительностью $\tau = 4$ час. за счет информационных атак на систему вышло из строя $M = 2$ БД. Тогда параметр λ , вычисленный по приведенной выше формуле, будет равен:

$$\lambda = 0,5 \text{ час}^{-1}.$$

Ожидаемое (среднее) число БД, которые могут выйти из строя в течение планируемой работы определяется как:

$$m_{\Delta N} = 0,5 \cdot 8 = 4 \text{ БД.}$$

Перейдем к расчету числа резервных копий БД, обеспечивающих успешное выполнение ПИТ с вероятностью $Q_{\text{дост}}$, не меньшей, чем значение:

$$P_{\text{тр}} = 0,95.$$

Параметр a , вычисленный по формуле (134), будет равен:

$$a = 46,87.$$

Табулируя значения функции (133), имеем:

$$\Delta N = 0, \quad \varphi(0) = 1 < a;$$

$$\Delta N = 1, \quad \varphi(1) = 5 < a;$$

$$\Delta N = 2, \quad \varphi(2) = 13 < a;$$

$$\Delta N = 3, \quad \varphi(3) = 23,667 < a;$$

$$\Delta N = 4, \quad \varphi(4) = 34,333 < a;$$

$$\Delta N = 5, \quad \varphi(5) = 42,867 < a;$$

$$\Delta N = 6, \quad \varphi(6) = 48,556 > a.$$

Отсюда делается вывод о том, что число резервных копий БД в планируемой работе должно удовлетворять условию:

$$\Delta N \geq 6 \text{ ед.}$$

Из приведенных результатов следует, что значение параметра $m_{\Delta N}$ может рассматриваться как приближенная оценка числа резервных копий БД.

Применим приведенную выше формулу для расчета ΔN из работы [Мамиконов А.Г. и др.]. Из теории надежности [Вентцель Е.С.] известно, что при заданном значении параметра λ вероятность разрушения БД в течение одного часа вычисляется как:

$$q = \int_0^1 \lambda e^{-\lambda t} dt = 1 - e^{-\lambda \cdot 1} = 1 - e^{-0,5 \cdot 1} = 0,393.$$

Тогда число резервных копий БД будет равно:

$$\Delta N = \left\lceil \frac{\ln(1 - 0,95)}{\ln 0,393} \right\rceil = [3,2076] = 4 \text{ ед.}$$

Отметим, что полученный в предложенном методе результат является для использованных в примере исходных данных несколько завышенным, но это только повышает надежность эксплуатации ПИТ.

Модель эксплуатации сервера ИС

При моделировании этого процесса будем использовать подход, изложенный в работе [Вентцель Е.С.]. Согласно ему процесс эксплуатации некоторой ИС описывается ветвящимся циклическим марковским процессом с конечным множеством состояний и непрерывным временем. Граф связи состояний сервера представлен на Рис. 70.

На этом рисунке приведены следующие состояния марковского процесса:

E_1 – сервер исправен и работает;

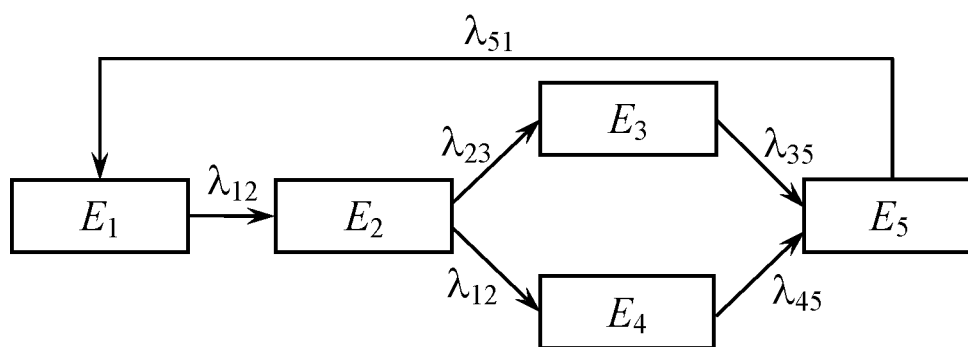


Рис. 70

E_2 – отказ сервера, диагностика неисправности;

E_3 – ликвидация неисправности техническим персоналом ИС;

E_4 – устранение неисправности специалистами центра сервисного обслуживания средств вычислительной техники;

E_5 – тестирование и запуск сервера.

Будем считать, что все потоки событий, переводящих процесс из состояния E_i в состояние E_j , являются пуассоновскими потоками, которые описываются интенсивностями λ_{ij} , $i, j \in \overline{1, 5}$, приведенными на Рис. 70.

Согласно этому предположению интенсивности таких потоков определяются следующим образом [Вентцель Е.С.]:

$$\begin{aligned} \lambda_{12} &= \frac{1}{T_1}; & \lambda_{23} &= \frac{q}{T_2}; & \lambda_{24} &= \frac{1-q}{T_2}; \\ \lambda_{35} &= \frac{1}{T_3}; & \lambda_{45} &= \frac{1}{T_4}; & \lambda_{51} &= \frac{1}{T_5}, \end{aligned}$$

где T_i – среднее время пребывания процесса в состоянии E_i , $i = \overline{1, 5}$; q – вероятность того, что неисправность сервера будет устранена техническим персоналом ИС.

Статическая вероятностная модель в форме 6-ти уравнений с 5-ю неизвестными приведена в указанной работе [Вентцель Е.С.]. Решение этой системы имеет вид:

$$\begin{aligned}
p_1 &= \frac{T_1}{T_1 + T_2 + qT_3 + (1-q)T_4 + T_5}; \\
p_2 &= \frac{T_2}{T_1 + T_2 + qT_3 + (1-q)T_4 + T_5}; \\
p_3 &= \frac{qT_3}{T_1 + T_2 + qT_3 + (1-q)T_4 + T_5}; \\
p_4 &= \frac{(1-q)T_4}{T_1 + T_2 + qT_3 + (1-q)T_4 + T_5}; \\
p_5 &= \frac{T_5}{T_1 + T_2 + qT_3 + (1-q)T_4 + T_5}.
\end{aligned} \tag{136}$$

Вероятности p_i определяют долю времени эксплуатации сервера, когда он находится в состоянии E_i , $i = (\overline{1, 5})$.

Пусть τ – общее время эксплуатации сервера в ИС. Тогда среднее время его пребывания в каждом из рассмотренных состояний будет равно:

$$\tau_i = p_i \tau, \quad i = (\overline{1, 5}). \tag{137}$$

Отсюда затраты на оплату специалистов центра сервисного обслуживания определяется как:

$$C = c\tau_4 = \frac{c(1-q)T_4}{T_1 + T_2 + qT_3 + (1-q)T_4 + T_5}, \tag{138}$$

где c – стоимость 1 часа работы привлеченных к ремонту специалистов.

Пример.

Пусть в некоторой ИС эксплуатируется сервер баз данных со следующими условными характеристиками его эксплуатации:

$$\begin{aligned}
T_1 &= 10000 \text{ ч}; \quad T_2 = 1 \text{ ч}; \quad T_3 = 4 \text{ ч}; \\
T_4 &= 24 \text{ ч}; \quad T_5 = 1 \text{ ч}; \quad q = 0,8; \quad c = 8000 \text{ руб./ч.}
\end{aligned}$$

Вероятности (136) состояния сервера будут равны:

$$p_1 = 0,9990; \quad p_2 = 0,000099; \quad p_3 = 0,00032;$$

$$p_4 = 0,000480; \quad p_5 = 0,000099;$$

Будем считать, что в течение года сервер эксплуатируется круглосуточно, т.е. $\tau = 365 \text{ дн.} \times 24 \text{ ч.} = 8760 \text{ ч.}$ Тогда из выражений (137) и (138) имеем:

$$\tau_1 = 8751,124 \text{ ч.;} \quad \tau_2 = 0,86 \text{ ч.} = 52 \text{ мин;}$$

$$\tau_3 = 2,8 \text{ ч.} = 168,2 \text{ мин;} \quad \tau_4 = 4,2 \text{ ч.} = 252 \text{ мин;}$$

$$\tau_5 = 0,86 \text{ ч.} = 52 \text{ мин;} \quad C = 33600 \text{ руб.}$$

Отметим, что по классификации математических моделей, приведенной в Лекции №4, предложенная модель является аналитически-разрешимой в форме выражений (136).

Самостоятельная работа.

Записать динамическую модель эксплуатации сервера в форме уравнений Колмогорова. Получить на ее основе статическую модель его эксплуатации в установившемся режиме. Решить полученные уравнения и получить формулы вида (136).

Прогнозирование количества функционирующих аппаратно-программных средств КИС

Пусть в некоторой КИС на момент времени t_0 функционирует N_i аппаратно-программных средств i -го вида, $i = (\overline{1, n})$. К таким средствам относятся АРМ персонала системы, серверы, коммутаторы, маршрутизаторы и др. Различные виды и структуры комплексов технических средств (КИС) приведены в монографиях [Зиновьев П.А., Мейко А.В., Моисеев В.С.].

Будем считать, что в процессе их эксплуатации каждый экземпляр этих средств может находиться в одном из следующих состояний марковского процесса эксплуатации аппа-

ратно-программных средств КИС: E_1 – исправен; E_2 – неисправен и диагностируется; E_3 – списан; E_4 – ремонтируется.

Граф связи этих состояний приведен на Рис. 71.

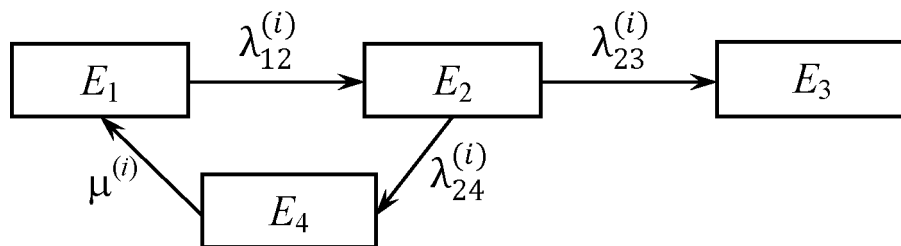


Рис. 71

На этом рисунке приведены интенсивности $\lambda_{jr}^{(i)}$, $j, r \in (\overline{1, 4})$ перехода процессов эксплуатации i -го средства КТС КИС между указанными состояниями; $\mu^{(i)}$ – интенсивность восстановления работоспособности каждого i -го средства, $i = (\overline{1, n})$.

Введем в рассмотрение вероятности $p^{(i)}$ того, что после диагностирования i -е средство будет списано, т.е. процесс его эксплуатации перейдет в поглощающее состояние E_3 . При этом с вероятностью $(1 - p^{(i)})$ оно будет отремонтировано (состояние E_4).

Будем считать, что потоки событий, переводящие процессы эксплуатации КТС КИС из состояний E_j в состояния E_r являются пуассоновскими потоками с указанными выше интенсивностями [Вентцель Е.С.].

Пусть $T_j^{(i)}$ – среднее время пребывания i -го средства в состоянии E_j , $j = (\overline{1, 4})$, $i = (\overline{1, n})$.

Тогда приведенные на Рис. 71 интенсивности определяются как [Вентцель Е.С.]:

$$\lambda_{12}^{(i)} = \frac{1}{T_1^{(i)}}; \quad \lambda_{23}^{(i)} = \frac{p^{(i)}}{T_2^{(i)}}; \quad \lambda_{24}^{(i)} = \frac{1 - p^{(i)}}{T_2^{(i)}};$$

$$\mu^{(i)} = \frac{1}{T_4^{(i)}}, \quad i = (\overline{1, n}).$$

За счет действия различных неопределенных факторов количество i -ых средств, находящихся в момент времени t в состоянии E_j , будут случайными функциями $X_j^{(i)}(t)$, $j = (\overline{1, 4})$, $i = (\overline{1, n})$.

Для представления таких функции в работе [Вентцель Е.С.] предлагается использовать метод «динамики средних», который позволяет вычислять такие характеристики исследуемых марковских случайных процессов, как их математическое ожидание и дисперсия.

Следуя этому методу, введем следующие обозначения:

$$M[X_j^{(i)}(t)] = m_j^{(i)}(t); \quad D[X_j^{(i)}(t)] = D_j^{(i)}(t), \quad i = (\overline{1, n}), \quad j = (\overline{1, 4}).$$

В указанной работе определение математических ожиданий осуществляется путем решения специальной системы дифференциальных уравнений, а дисперсии строятся как функции от их значений.

В нашем случае для описания процессов эксплуатации КТС КИС на заданном интервале времени $[t_0, t_k]$ будут использованы n систем таких уравнений, описывающих динамику изменения среднего количества i -х средств, находящихся в состояниях E_1, E_2, E_3, E_4 (см. Рис. 71):

$$\begin{aligned} \dot{m}_1^{(i)} &= \frac{m_4^{(i)}}{T_4^{(i)}} - \frac{m_1^{(i)}}{T_1^{(i)}}; & \dot{m}_2^{(i)} &= \frac{m_1^{(i)}}{T_1^{(i)}} - \frac{m_2^{(i)}}{T_2^{(i)}}; \\ \dot{m}_3^{(i)} &= \frac{p^{(i)} m_2^{(i)}}{T_2^{(i)}}; & \dot{m}_4^{(i)} &= \frac{(1-p^{(i)}) m_2^{(i)}}{T_2^{(i)}} - \frac{m_4^{(i)}}{T_4^{(i)}}, \end{aligned} \quad (139)$$

$$i = (\overline{1, n}), \quad t \in [t_0, t_k].$$

Дисперсии количества i -ых средств, находящихся в момент времени t в состоянии E_j , предлагается вычислять по формуле:

$$D_j^{(i)}(t) = m_j^{(i)}(t) \left(1 - \frac{m_j^{(i)}(t)}{N_i} \right), \quad (140)$$

$$t \in [t_0, t_k], \quad i = (\overline{1, n}), \quad j = (\overline{1, 4}).$$

Значения функции $m_j^{(i)}(t)$, определенные из решения систем уравнений (139), должны удовлетворять равенствам:

$$m_1^{(i)}(t) + m_2^{(i)}(t) + m_3^{(i)}(t) + m_4^{(i)}(t) = N_i, \quad (141)$$

$$t \in [t_0, t_k], \quad i = (\overline{1, n}).$$

При этом считается, что в начальный момент времени $t = t_0$ все средства КИС работоспособны. Тогда начальные условия для интегрирования систем (139) примут вид:

$$m_1^{(i)}(t_0) = N_i; \quad m_j^{(i)}(t_0) = 0, \quad j = (\overline{1, 4}).$$

Полученная модель эксплуатации i -го средства является избыточной, включающей в себя 5 уравнений (139), (141), содержащих 4 неизвестные функции $m_1^{(i)}(t)$, $m_2^{(i)}(t)$, $m_3^{(i)}(t)$, $m_4^{(i)}(t)$. Поэтому для выполнения условия (141) выделяют одну из них, например:

$$m_1^{(i)}(t) = N_i - \sum_{j=2}^4 m_j^{(i)}(t), \quad i = (\overline{1, n}) \quad (142)$$

и в системе (139) отбрасывают первое уравнение, а в оставшихся проводят соответствующую замену.

Тогда преобразованная задача Коши запишется как:

$$\begin{aligned}
\dot{m}_2^{(i)} &= \frac{1}{T_1} \left(N_i - \sum_{j=2}^4 m_j^{(i)} \right) - \frac{m_2^{(i)}}{T_2}; \\
\dot{m}_3^{(i)} &= \frac{p^{(i)} m_2^{(i)}}{T_2}; \\
\dot{m}_4^{(i)} &= \frac{(1 - p^{(i)}) m_2^{(i)}}{T_2} - \frac{m_4^{(i)}}{T_4}, \quad i = (\overline{1, n}), \quad t \in [t_0, t_k]; \\
m_2^{(i)}(t_0) &= 0; \quad m_3^{(i)}(t_0) = 0; \quad m_4^{(i)}(t_0) = 0.
\end{aligned} \tag{143}$$

Определение среднего числа работоспособных в момент времени $t \in [t_0, t_k]$ средств i -го вида осуществляется путем подстановки результатов ее решения в формулу (142). При этом дисперсии числа средств, находящихся в текущий момент времени t в различных состояниях, вычисляются по формуле (140).

Для более достоверной оценки прогнозируемого состояния КТС КИС предлагается применить известное в теории вероятностей «правило 3σ ». Это правило применительно к решаемой задаче можно представить как:

$$P\{m_j^{(i)}(t) - 3\sigma_j^{(i)}(t) \leq X_j^{(i)}(t) \leq m_j^{(i)}(t) + 3\sigma_j^{(i)}(t)\} \approx 1,$$

где

$$\sigma_j^{(i)}(t) = \sqrt{D_j^{(i)}(t)} = \sqrt{m_j^{(i)}(t) \left(1 - \frac{m_j^{(i)}(t)}{N_i} \right)}, \quad i = (\overline{1, n}), \quad j = (\overline{1, 4}).$$

При этом для прогнозирования количества заменяемых списанных средств новыми средствами будем использовать соотношение:

$$N_i^*(t) = \left[m_4^{(i)}(t) + 3 \sqrt{m_4^{(i)}(t) \left(1 - \frac{m_4^{(i)}(t)}{N_i} \right)} \right], \quad t \in [t_0, t_k]. \tag{144}$$

где $[(\cdot)]$ – операция округления значений (\cdot) до ближайших целых значений в большую сторону.

Рассмотренная в этой задаче математическая модель является непрерывной аналитической динамической вероятностной численно-разрешимой моделью (см. Рис. 13).

Пример.

Пусть в составе КТС КИС крупного машиностроительного объединения имеется $N_1 = 1200$ АРМ, $N_2 = 10$ серверов и $N_3 = 40$ коммутаторов [Зиновьев П.А., Мейко А.В., Моисеев В.С.].

По данным этой монографии среднее время их наработки на отказ соответственно равны:

$$T_1^{(1)} = 10000 \text{ ч}; \quad T_1^{(2)} = 50000 \text{ ч}; \quad T_1^{(3)} = 200000 \text{ ч}.$$

Вероятности списания этих средств имеют следующие значения:

$$p^{(1)} = 0,1; \quad p^{(2)} = 0,05; \quad p^{(3)} = 0,5.$$

Средние затраты времени на диагностирование АРМ, серверов и коммутаторов равны:

$$T_2^{(1)} = 4 \text{ ч}; \quad T_2^{(2)} = 10 \text{ ч}; \quad T_2^{(3)} = 2 \text{ ч}.$$

Среднее время на их ремонт:

$$T_4^{(1)} = 33,33 \text{ ч}; \quad T_4^{(2)} = 100 \text{ ч}; \quad T_4^{(3)} = 20 \text{ ч}.$$

Используем эти исходные данные для расчета динамики изменения числа списываемых (заменяемых) в течение года АРМ, серверов и коммутаторов.

Применяя при их круглосуточной работе выражения (142)-(144) на интервале времени $[0; 8760]$ ч., получаем ежеквартальные результаты, представленные в Табл. 6.

Эти результаты взяты из Табл. ПЗ.1 указанной выше книги с разбивкой по кварталам и следующими обозначениями:

$$N_1^* = N_{\text{АРМ}}^*; \quad N_2^* = N_{\text{серв.}}^*; \quad N_3^* = N_{\text{комм.}}^*.$$

Таблица 6

квартал	I	II	III	IV
t , ч.	2208	4406	6605	8760
$N_{\text{АРМ}}^*$	36	63	88	112
$N_{\text{серв.}}^*$	1	1	1	1
$N_{\text{КОММ.}}^*$	2	3	4	4

Из Табл. 6 следует, что в течение года функционирования КИС возможна замена 1 сервера, 4 коммутаторов и 112 АРМ персонала системы.

Кроме этого результаты прогнозирования дают оценки потребного количества «запасных» аппаратно-программных средств системы.

Самостоятельная работа.

Изучить метод «динамики средних» и найти его применение в задачах прикладной информатики.

Модель динамики количества пользователей в КИС

Рассмотренное в предыдущих лекциях понятие корпоративной информационной системы (КИС) нуждается в дополнительных исследованиях. В частности, в существующей литературе практически отсутствуют задачи управления развитием этого широко распространенного вида информационных систем. Поэтому кратко изложим суть таких задач. Тем более, что рассматриваемая модель является одной из них.

При совершенствовании процессов эксплуатации и разработки стратегии КИС представляется необходимым иметь адекватный математический аппарат для мониторинга и прогнозирования динамики развития задействованных в ней системных ресурсов. При этом мониторингом должны быть охвачены все основные характеристики организационно-технического состояния исследуемой системы.

Под описанием организационно-технического состояния КИС в определенный момент времени будем понимать значения следующих функциональных показателей [Гиннатуллин И.А., Зиновьев П.А., Моисеев В.С.]:

- число активных пользователей;
- перечень и количество решаемых ими задач с указанием используемых комплексов прикладных программ;
- совокупность и количество используемых баз данных (БД) распределенного банка данных (РБнД) системы;
- перечень и общее количество активных аппаратно-программных средств (концентраторов, маршрутизаторов, серверов, АРМ и т.п.);
- перечень и общее количество активных системных программных средств;
- плотность трафиков в каналах связи системы.

Мониторинг состояния системы предусматривает отслеживание, фиксацию и оценку значений этих показателей с определенной периодичностью. Получаемые при этом результаты используются в задачах управления системными ресурсами и процессами.

Управление КИС осуществляется для достижения следующих основных целей:

1. Обеспечение оптимального уровня производительности системы;
2. Обеспечение требуемого качества обслуживания пользователей и приложений;
3. Эффективное использование корпоративных информационно-вычислительных ресурсов.

Математическая модель для выбора оптимальной периодичности контроля (мониторинга) состояния КИС должна в

качестве минимизируемого критерия учитывать общий объем используемых временных и материальных ресурсов. При этом для различных компонентов КИС могут выбираться собственные величины периодов контроля их состояний. В модели должны быть также учтены реальные характеристики надежности технических и программных средств КИС и потоков обращений к системе различных категорий пользователей или подразделений организации. Для получения последних в программном и информационном обеспечении мониторинга должны быть предусмотрены средства сбора, хранения и статистической обработки соответствующих данных, получаемых на различных интервалах функционирования КИС в режиме «24 × 365».

Методы управления текущим состоянием КИС используются при решении следующих задач:

- расчет значений фактической и прогнозируемой производительности КИС;
- определение фактической достоверности результатов работы системы;
- выделение сегментов КИС, в которых нарушаются заданные значения производительности и/или достоверности;
- диагностика технического состояния используемых в каждом сегменте КИС технических и программных средств, а также фрагментов применяемых БД;
- принятие решений по маршрутизации информационных потоков, по распараллеливанию процессов выполнения прикладных информационных технологий (ПИТ), замене технических средств или подключению дополнительных устройств и т.п.

Задача диагностики может быть решена путем выделения в сегменте КИС отдельных маршрутов передачи обрабатываемых или запрашиваемых данных и реализации процедуры выявления элементов, отрицательно влияющих на производительность системы и/или достоверность результатов.

При построении предлагаемых математических моделей необходимо использовать модели развивающихся технико-экономических систем [Кучин Б.Л., Якушенко Е.В.], к которым относится КИС.

Вследствие относительно небольшого числа пользователей и объемов используемых в составе КИС данных (по сравнению с приведенными в этой работе в моделях Риденура и Гартмана числом пользователей перспективных технологий и объемом циркулирующей в мире научно-технической информации) преобразуем их непрерывные модели, представленные в виде дифференциальных уравнений, в конечно-разностные уравнения на дискретной шкале времени функционирования КИС.

Пусть L_t – число пользователей ПИТ, реализованных в составе КИС в момент времени t ; L_{\max} – верхний предел для количества пользователей рассматриваемой КИС.

Рассмотрим момент времени $t+1$ и будем считать, что приращение числа пользователей в интервале $(t, t+1)$ будет пропорционально числу пользователей системы, знакомых с содержащейся в ней информацией, и числу пользователей, которые в момент времени t не используют в своей работе содержащуюся в системе информацию.

Формализуем эту гипотезу в виде следующего равенства:

$$\Delta L_{t+1} = L_{t+1} - L_t = aL_t(L_{\max} - L_t),$$

где $a > 0$ – коэффициент пропорциональности.

Перепишем это соотношение в виде нелинейного разностного уравнения:

$$L_{t+1} = L_t + aL_t(L_{\max} - L_t), \quad t = 0, 1, 2, \dots \quad (145)$$

Для расчета динамики изменения числа пользователей с помощью этого уравнения необходимо использовать начальное условие вида:

$$L_0 = L^0, \quad (146)$$

где L^0 – фактическое число пользователей КИС в момент времени $t = 0$.

В процессе реального функционирования КИС число ее потенциальных пользователей L_{\max} может изменяться за счет сокращения или расширения числа подразделений соответствующей организации. В этом случае имеем функцию $L_{\max} = L_{\max}(t)$, которая на определенных интервалах времени $(0, \tau_1)$, $(\tau_1, \tau_1 + \tau_2), \dots$ реорганизации корпорации имеет вид кусочно-постоянной функции.

Построим модель динамики изменения объемов информации, хранящихся в КИС. Предполагается, что пополнение и исключение информации осуществляются только пользователями данной системы.

Будем считать, что приращение объема информации в РБнД КИС пропорциональны числу пользователей системы в фиксированный момент времени, а также объему уже хранимой информации, используемой пользователями, но не представленной в базах данных системы.

Пусть J_t, J_{t+1} – объем хранимой информации в моменты времени t и $t + 1$; L_t – число пользователей системы в момент времени t .

Тогда сформулированную гипотезу можно записать формально в виде соотношения:

$$\Delta J_{t+1} = J_{t+1} - J_t = b L_t J_t (J_{\max} - J_t),$$

где $b > 0$ – коэффициент пропорциональности, J_{\max} – верхний предел объема информации, используемого системой, который также может быть переменным, т.е. $J_{\max} = J_{\max}(t)$.

Перепишем это соотношение в виде нелинейного разностного уравнения:

$$J_{t+1} = J_t + b L_t J_t (J_{\max} - J_t), \quad t = 0, 1, 2, \dots \quad (147)$$

с начальным условием вида:

$$J_0 = J^0, \quad (148)$$

где J^0 – объем данных, хранящихся в системе в момент времени $t = 0$.

Таким образом, для анализа динамики числа пользователей и связанных с ними объемов хранимой в РБнД КИС информации необходимо решить систему разностных уравнений (145), (147) с начальными условиями (146), (148).

Эта модель, которая является дискретной динамической детерминированной моделью, может быть использована для прогнозирования этих важнейших показателей работы КИС, но только в случае указания конкретных числовых значений коэффициентов a и b , входящих в уравнения (145), (147). Для каждой конкретной организации эти коэффициенты могут быть определены путем сбора и обработки фактических значений показателей \bar{L}_t и \bar{J}_t , на некотором интервале времени $[\theta_0, \theta_k]$. Определим параметр a с помощью известного метода наименьших квадратов.

Перепишем уравнение (145) в виде:

$$L_{t+1} - L_t - a L_t (L_{\max} - L_t) = 0,$$

На его основе построим следующую функцию невязки:

$$\delta = \sum_{t=0}^k [\bar{L}_{t+1} - \bar{L}_t - a \bar{L}_t (L_{\max} - \bar{L}_t)]^2.$$

Коэффициент a будем определять из условия минимума функции $\delta = \delta(a)$.

Используя необходимые условия экстремума этой функции, имеем:

$$\frac{d\delta}{da} = -2 \sum_{t=0}^k [(\bar{L}_{t+1} - \bar{L}_t - a\bar{L}_t(L_{\max} - \bar{L}_t))\bar{L}_t(L_{\max} - \bar{L}_t)] = 0.$$

Откуда значение a вычисляется как:

$$a = \frac{\sum_{t=0}^k [(\bar{L}_{t+1} - \bar{L}_t)\bar{L}_t(L_{\max} - \bar{L}_t)]}{\sum_{t=0}^k [\bar{L}_t^2(L_{\max} - \bar{L}_t)]}. \quad (149)$$

Аналогичным образом может быть получена расчетная формула для определения коэффициента b , входящего в уравнение (147).

Пример.

Выполним прогнозирование числа почтово-кассовых терминалов (ПКТ) в процессе информатизации почтовой отрасли г. Казани, которая охватывает практически все городские подразделения почтовой связи. Поэтому очень важно иметь долгосрочные и среднесрочные прогнозы числа операторов-кассиров, работающих с ПКТ КИС. Эта информация необходима для выявления направлений развития локальных информационных систем, каналов связи и т.п. Оценив перспективы развития, можно определить будущие нагрузки на узлы КИС почтовой связи г. Казани, заблаговременно выявить возникновение «узких» мест по переработке информации и т.д.

Во втором столбце Табл. 7 приведены фактические данные о динамике числа ПКТ в г. Казани в период с 1996 по 2002 год.

Таблица 7

t , год	\bar{L}_t , ед.	L_t , ед.	ΔL , ед.	ΔL , %
1996	11	11	0	0
1997	116	116	0	0
1998	178	160	18	10,1
1999	247	217	30	12,4
2000	300	287	13	4,3
2001	458	367	91	19,8
2002	460	451	9	1,9
2003	—	533	—	—
2004	—	602	—	—
2005	—	656	—	—
2006	—	693	—	—
2007	—	717	—	—
2008	—	731	—	—

Используя модель (145), (146), сформируем пятилетний прогноз по внедрению новых ПКТ в КИС почтовой связи г. Казани.

Для применяемой модели нужно определить максимальное число ПКТ в системе. При выборе значения L_{\max} можно ориентироваться на следующие данные: в г. Казани находится 150 почтовых отделений, в каждом из которых может разместиться не более 5 ПКТ. Таким образом, получим $L_{\max} = 750$ ПКТ.

По формуле (149), используя значения \bar{L}_t , найдем значение коэффициента a в уравнении (145):

$$a = 0,000603.$$

Проводя вычисления по формуле (145) при начальном условии (146) вида $L_{1996} = 11$, получаем прогнозируемую зависимость L_t числа ПКТ до 2008 года, представленные в третьем столбце Табл. 7.

В четвертом и пятом столбцах этой таблицы даны значения отклонений фактических от прогнозируемых данных. Из их анализа следует, что максимальная ошибка прогнозирования составляет величину порядка 19%, а ее среднее значение равно 7%. Такие значения являются приемлемыми при использовании приведенных выше простейших моделей.

Из результатов расчетов следует, что к 2008 году число пользователей КИС должно увеличиться до 731 операторов-кассиров. Необходимо заметить, что для рассматриваемой модели очень важно оценить максимальное возможное число терминалов. Данная оценка может меняться с течением времени. В этом случае необходимо рассчитать параметр модели a заново.

В заключение лекции отметим, что приведенные в ней модели, несомненно, нуждаются в дальнейшем развитии.

Раздел VI. МОДЕЛИ СУЩЕСТВУЮЩИХ И ПЕРСПЕКТИВНЫХ ТЕХНОЛОГИЙ ЗАЩИТЫ ИНФОРМАЦИИ

Проблема обеспечения информационной безопасности является ключевой в современном компьютеризированном обществе. Практика успехов киберпреступности показала полный провал существующих подходов к защите информации с их псевдонаучной терминологией («стратегии безопасности», «профили безопасности», «красный, желтый и еще какие-то уровни опасности в системе» и прочая и прочая) и «мутными» методиками ее обеспечения.

На наш взгляд, успехов в решении этой проблемы можно добиться только при создании строгой прикладной теории информационной безопасности, базирующейся на адекватных проверенных практикой математических моделях и четких методиках создания и применения перспективных программных и аппаратных средств защиты информации.

21. Модель традиционной системы защиты информации

Пусть несанкционированному доступу [Мамиконов А.Г., Кульба В.В., Шелков А.Б.] могут быть подвергнуты следующие компоненты информационной системы (ИС):

A – множество автоматизированных рабочих мест и других аппаратных средств системы;

P – множество системных и прикладных программ;

Φ – множество файлов системы.

Обозначим через $U = \{u_1, u_2, \dots, u_n\}$ множество действующих угроз рассматриваемой системе. В состав этого множества могут входить следующие известные на сегодняшний день действия ее противника:

u_1 – перехват электромагнитных излучений,

u_2 – принудительное облучение (подсветка) линий связи с целью внесения помех,

u_3 – применение подслушивающих устройств,

u_4 – дистанционное фотографирование,

u_5 – перехват акустических колебаний и восстановление текста, печатаемого акустическими принтерами,

u_6 – хищение носителей информации,

u_7 – считывание данных в файлах пользователей,

.....

u_{17} – использование компьютерных вирусов.

Для исключения последствий угроз используется множество $M = \{m_1, m_2, \dots, m_k\}$ средств защиты информации, которое в общем случае можно разбить на следующие подмножества: M_1 – аппаратные средства, M_2 – физические средства (ограждения, решётки), M_3 – программные средства, M_4 – организационные средства, M_5 – законодательные средства, M_6 – морально-этические средства.

Пусть множество объектов защиты имеет вид:

$$Z = A \cup P \cup \Phi. \quad (150)$$

Зададим первичные отношения вида «угроза-объект» и «объект-средство защиты» в виде:

$$R_1 \subseteq U \times Z; \quad R_2 \subseteq Z \times M. \quad (151)$$

Тогда простейшая бинарная модель системы защиты информации (СЗИ) задаётся формальным объектом вида:

$$\mathfrak{M}_{\text{сзи}} = \{Z, R_1, R_2\}.$$

Взаимосвязь угроз и средств защиты можно определить с помощью вторичного отношения вида:

$$Q_1 = R_1 \circ R_2 \subseteq U \times M.$$

Это отношение представим в виде матрицы $B_1 = [b_{ij}]$, где $b_{ij} = 1$, если от i -й угрозы защищает j -е средство и $b_{ij} = 0$ в противном случае.

Построение этой матрицы позволяет формальным образом провести первичный анализ СЗИ. В частности, наличие в матрице Q_1 нулевой строки с номером s говорит о том, что для угрозы $u_s \in U$ не существует средств защиты.

Тернарное отношение:

$$Q_2 = Q_1 \cup R_2^{-1} \subseteq U \times M \times Z.$$

позволяет построить «трёхслойный» ориентированный граф и проследить направление атак и существующие средства защиты от них для каждого конкретного компонента ИС.

Пример.

Пусть защищаемыми компонентами некоторой системы являются следующие множества:

$$A = \{a_1, a_2\}, P = \{p_1, \dots, p_4\}, \Phi = \{f_1, f_2, f_3\}.$$

Будем считать, что множество U включает в себя следующие действия противника:

u_1 — несанкционированное копирование,

u_2 — доступ к закрытым файлам,

u_3 — доступ к АРМ.

Множество M включает в себя следующие мероприятия:

m_1 — охрана помещений,

m_2 — защита с помощью паролей,

m_3 — криптографическая защита,

m_4 — защита от электромагнитного излучения аппаратных средств,

m_5 — контроль целостности программ и данных.

Множество Z записывается следующим образом:

$$Z = \{a_1, a_2, p_1, \dots, p_4, f_1, \dots, f_3\}$$

Отношения вида (151) конкретизируются в матричной форме как:

$$R_1 = \begin{matrix} & a_1 & a_2 & p_1 & p_2 & p_3 & p_4 & f_1 & f_2 & f_3 \\ \begin{matrix} u_1 \\ u_2 \\ u_3 \end{matrix} & \left| \begin{array}{ccccccccc} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right. \end{matrix};$$

$$R_2 = \begin{matrix} & m_1 & m_2 & m_3 & m_4 & m_5 \\ \begin{matrix} a_1 \\ a_2 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ f_1 \\ f_2 \\ f_3 \end{matrix} & \left| \begin{array}{ccccc} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array} \right. \end{matrix}.$$

Сформируем модель работы пользователей с защищенными файлами ИС. Пусть при работе ИС множество пользователей Π может выполнять с множеством файлов Φ следующие действия:

- ЧТЕНИЕ содержания конкретных файлов
- ИЗМЕНЕНИЕ содержимого определённых файлов.

Возможности пользователей при выполнении этих действий опишем отношениями вида:

$$R_1 \subseteq \Pi \times \Phi; \quad R_2 \subseteq \Pi \times \Phi. \quad (152)$$

Пример.

Пусть $\Pi = \{\pi_1, \pi_2, \pi_3\}$, $\Phi = \{f_1, f_2, f_3\}$.

Для этих множеств отношения (152) конкретизируется как:

$$R_1 = \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \end{array} \left| \begin{array}{ccc} f_1 & f_2 & f_3 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{array} \right|; \quad R_2 = \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \end{array} \left| \begin{array}{ccc} f_1 & f_2 & f_3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{array} \right|.$$

Выявление пользователей, которые могут и читать, и изменять содержимое файлов может быть осуществлено с помощью вторичного отношения вида:

$$Q_1 = R_1 \cap R_2.$$

В матричной форме это отношение получается как (см. Лекцию №6):

$$Q_1 = R_1 \otimes R_2.$$

В условиях этого примера матричное представление отношения Q_1 примет следующий вид:

$$Q_1 = \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \end{array} \left| \begin{array}{ccc} f_1 & f_2 & f_3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right|.$$

Пусть в рассматриваемой ИС имеется три традиционных уровня конфиденциальности данных, описываемых множеством:

$$Y = \{y_1, y_2, y_3\},$$

где y_1 – уровень «Для служебного пользования»; y_2 – уровень «Секретно»; y_3 – уровень «Совершенно секретно».

Допуск пользователей системы к соответствующей информации будем описывать отношением вида:

$$R_3 \subseteq \Pi \times Y.$$

Соответственно, уровень конфиденциальности информации, хранимой в файлах системы, описывается как отношение:

$$R_4 \subseteq \Phi \times Y.$$

Тогда возможности работы пользователей ИС с соответствующими файлами представляется вторичным отношением:

$$Q_2 = R_3 \circ R_4^{-1} \subseteq \Pi \times \Phi.$$

Пусть для рассмотренного выше примера отношения R_3 и R_4 конкретизируются как:

$$R_3 = \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \end{array} \left| \begin{array}{ccc} y_1 & y_2 & y_3 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{array} \right| ; \quad R_4 = \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \end{array} \left| \begin{array}{ccc} y_1 & y_2 & y_3 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{array} \right|.$$

Тогда отношение Q_2 принимает следующий вид:

$$Q_2 = \begin{array}{c} \pi_1 \\ \pi_2 \\ \pi_3 \end{array} \left| \begin{array}{ccc} f_1 & f_2 & f_3 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right|.$$

Из этой матрицы следует, что пользователь π_1 имеет доступ к файлам f_2 и f_3 , пользователь π_2 — ко всем файлам системы, а пользователь π_3 не имеет права доступа к файлам f_1, f_2, f_3 .

Таким образом, модель работы пользователей ИС с конфиденциальной информацией имеет вид:

$$\mathfrak{M}_{\text{РКИ}} = \{\Pi, \Phi, Y, R_1, \dots, R_4\}.$$

Предложенные выше теоретико-множественные модели должны выступать в качестве исходных данных для решения задач анализа и синтеза ИС.

Рассмотрим простейшие задачи выбора оптимальных средств защиты ИС.

Задача №1

Пусть, используя приведенную выше модель $\mathfrak{M}_{\text{СИ}}$ системы защиты информации, и в частности отношение $Q_1 \subseteq U \times M$, которое описывается булевой матрицей $B_1 = [b_{ij}]$ с элемен-

тами $b_{ij} = 1$, если от угрозы $u_i \in U$ защищает средство $m_j \in M$, и $b_{ij} = 0$ в противном случае.

Требуется выбрать совокупность средств защиты из множества M , обеспечивающих их минимальную стоимость и минимум вероятности «взлома» системы.

Введём переменные:

$$x_j \in \{0; 1\}, \quad j = (\overline{1, m}), \quad (153)$$

такие, что $x_j = 1$, если для защиты объектов выбрано j -е средство, $x_j = 0$ – в противном случае.

Пронумеруем элементы множеств угроз и средств защиты следующим образом:

$$U = \{1, 2, \dots, n\}, \quad M = \{1, 2, \dots, m\}.$$

Условие того, что любая угроза будет отражена не менее чем одним средством защиты, записывается следующей системой неравенств:

$$\sum_{j=1}^m b_{ij} x_j \geq 1, \quad i = (\overline{1, n}). \quad (154)$$

Стоимость использованных средств защиты ИС определяется выражением:

$$C = \sum_{j=1}^m c_j x_j \rightarrow \min, \quad (155)$$

где c_j – стоимость реализации j -го средства защиты.

Пусть p_j – вероятность того, что противник сможет преодолеть j -е средство защиты информации. Практика показала, что чем больше c_j , тем меньше вероятность p_j .

Будем считать, что противник одновременно атакует все существующие в ИС средства защиты. В этом случае вероятность «взлома» выбранных средств защиты может быть определена как:

$$P = p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} = \prod_{j=1}^m p_j^{x_j} \rightarrow \min. \quad (156)$$

Отметим, что если все $x_j = 0$, то из этого выражения следует, что вероятность $P = 1$.

Таким образом, для выбора средств защиты ИС получили двухкритериальную задачу нелинейного булевского программирования вида (153)-(156).

Преобразуем критерий (156) к линейному виду, прологарифмировав обе его части:

$$\ln P = \sum_{j=1}^m x_j \ln p_j.$$

Вводя обозначения $W = \ln P$, $W_j = \ln p_j$, получим линейный критерий вида:

$$W = \sum_{j=1}^m W_j x_j. \quad (157)$$

В этом случае преобразованную задачу, которая является развитием задачи о наименьшем (минимальном) покрытии [Кристофидес Н.], можно решать численными методами, изложенными в работе [Сигал И.Х., Иванова А.В.].

Пример.

Пусть множество угроз ИС имеет вид $U = \{u_1, u_2, u_3\}$. Будем считать, что разработчик системы защиты ИС располагает множеством средств защиты $M = \{m_1, m_2, m_3, m_4\}$. Для этого множества имеют место следующие характеристики:

$$\begin{aligned} c_1 = 5; \quad p_1 = 0,3; \quad c_2 = 10, \quad p_2 = 0,1; \\ c_3 = 1, \quad p_3 = 0,8; \quad c_4 = 6, \quad p_4 = 0,2. \end{aligned}$$

Отношение Q_1 описывается матрицей:

$$B_1 = \begin{array}{c|cccc} & m_1 & m_2 & m_3 & m_4 \\ u_1 & 0 & 1 & 1 & 1 \\ u_2 & 0 & 1 & 0 & 1 \\ u_3 & 1 & 0 & 0 & 0 \end{array}.$$

Условия (153) конкретизируются как:

$$x_1 \in \{0, 1\}, x_2 \in \{0, 1\}, \dots, x_4 \in \{0, 1\}. \quad (158)$$

Система ограничений (154) записывается как:

$$\begin{aligned} x_2 + x_3 + x_4 &\geq 1; \\ x_2 + x_4 &\geq 1; \\ x_1 &\geq 1. \end{aligned} \quad (159)$$

Стоимость выбранных средств защиты (155) задается выражением:

$$C = 5x_1 + 10x_2 + 1x_3 + 6x_4 \rightarrow \min.$$

Целевая функция (156) имеет вид:

$$P = 0,3^{x_1} \cdot 0,1^{x_2} \cdot 0,8^{x_3} \cdot 0,2^{x_4} \rightarrow \min.$$

Нетрудно показать, что множество допустимых решений задачи, определяемых условиями (158) и (159), включает в себя две точки с координатами:

1) $x_1^0 = 1; x_2^0 = 1; x_3^0 = 0; x_4^0 = 0$ (используется 1-е и 2-е средства защиты).

2) $x_1^0 = 1; x_2^0 = 0; x_3^0 = 0; x_4^0 = 1$ (применяется 1-е и 4-е средства).

Значения целевых функций C и P в этих точках будут соответственно равны:

$$C_1 = 5 \cdot 1 + 10 \cdot 1 + 1 \cdot 0 + 6 \cdot 0 = 15;$$

$$P_1 = 0,3^1 \cdot 0,1^1 \cdot 0,8^0 \cdot 0,2^0 = 0,03;$$

$$C_2 = 11; P_2 = 0,06.$$

На Рис. 72 представлены эти значения в пространстве критериев решаемой задачи.

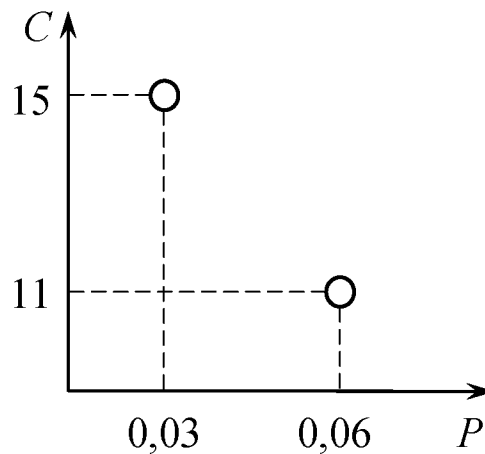


Рис. 72

Из этого множества оптимальных по Парето решений задачи Разработчик системы защиты ИС может выбрать конкретный компромиссный вариант между стоимостью СЗИ и вероятностью ее «взлома».

Задача №2

Пусть для любого j -го средства защиты имеется r_j вариантов его реализации, $j = (\overline{1, m})$. Будем считать, что любая k -я реализация j -го средства защиты позволяет с вероятностью q_{kji} «исключить» i -ю угрозу, $i = (\overline{1, n})$. Стоимость k -ой реализация j -го средства обозначим через C_{kj} .

Требуется для любого средства защиты выбрать одну реализацию так, чтобы стоимость СЗИ и вероятность её преодоления противником были минимальны.

Введем булевские переменные y_{kj} , причём $y_{kj} = 1$, если j -ое средство реализуется в СЗИ в k -ом варианте и $y_{kj} = 0$ – в противном случае.

Условие выбора одного варианта каждого средства имеет вид:

$$\sum_{k=1}^{r_j} y_{kj} = 1, \quad j = (\overline{1, m}). \quad (160)$$

Стоимость системы защиты определяется так:

$$C = \sum_{j=1}^m \sum_{k=1}^{r_j} C_{kj} y_{kj} \rightarrow \min. \quad (161)$$

Введём гипотезу H_i – «противник использует для вскрытия системы i -ю угрозу». Обозначим через $p(H_i) = \alpha_i$ вероятность применения i -ой угрозы.

Предположим, что противник будет применять только одну угрозу. Это означает, что $\sum_i \alpha_i = 1$.

Пусть A – случайное событие, состоящее в преодолении системы защиты. Вероятность преодоления защиты по формуле полной вероятности может быть определена как

$$P(A) = \sum_{i=1}^n p(A | H_i) p(H_i).$$

Определим условную вероятность $p(A | H_i)$ преодоления защиты i -й угрозой. Согласно матрице $B_1 = [b_{ij}]$ от этой угрозы защитой является подмножество средств $M_i = \{m_j \in M \mid b_{ij} = 1\}$. С учётом заданных значений q_{kji} и b_{ij} имеем, что:

$$p(A | H_i) = \sum_{m_j \in M_i} \sum_{k=1}^{r_j} b_{ij} (1 - q_{kji}) y_{kj}.$$

Тогда второй критерий задачи принимает вид:

$$P(A) = \sum_{i=1}^n \sum_{m_j \in M_i} \sum_{k=1}^{r_j} \alpha_i b_{ij} (1 - q_{kji}) y_{kj} \rightarrow \min. \quad (162)$$

Постановка задачи №2 конкретизируется как: «Найти значения переменных y_{kj} , принадлежащих множеству $\{0; 1\}$, доставляющих минимум по Парето целевым функциям (161) и (162) при удовлетворении условия (160)».

Самостоятельная работа.

Для решения приведенной двухкритериальной задачи булевского программирования разработать эвристический алгоритм, учитывающий ее ограничения вида (160).

Другие задачи оптимизации традиционных систем защиты информации приведены в монографии [Кузнецов Н.А., Кульба В.В., Микрин В.А.].

22. Теоретико-игровая вероятностная модель шифрования и дешифрования данных с применением алгоритмов DES и RSA

Алгоритмы DES и RSA, обладающие за счет их высокой стойкости достаточно широким распространением в различных информационных системах, практически с моментов времени их появления постоянно подвергаются попыткам их «взлома».

В данной лекции предлагаются методы их совместного рандомизированного (случайного) использования в разрабатываемых комплексах программ (КП), позволяющих повысить их стойкость к информационным атакам противника («хакеров»).

В основу метода положена математическая модель игры с нулевой суммой размерности (2×2) . В виде способов, конкретизирующих применение в КП алгоритмов DES или RSA, используются понятия физической смеси оптимальных смешанных стратегий игры и метода Монте-Карло [Вентцель Е.С.].

В качестве игроков будем рассматривать Разработчика КП и «хакера», которые имеют следующие стратегии:

A_1 – применение в КП модуля, реализующего алгоритм DES;

A_2 – применение в КП модуля, реализующего алгоритм RSA;

B_1 – использование «хакером» аппаратно-программных средств (АПС) «взлома» алгоритма DES;

B_2 – использование им АПС «взлома» алгоритма RSA.

Будем считать, что стойкость рассматриваемых алгоритмов оценивается затратами времени T_1 и T_2 на их «взлом».

Считается, что если «хакер» знает о применении в КП алгоритма DES (стратегия A_1), то он, используя стратегию B_1 , затрачивает на «взлом» КП T_1 ед. вр.

Если на стратегию A_1 «хакер» отвечает стратегией B_2 , то его затраты времени на «взлом» будут равны $(T_2 + T_1)$ ед. вр. Последнее означает, что, использовав без успеха АПС для «взлома» алгоритма RSA и поняв свою ошибку, «хакер» применяет АПС «взлома» алгоритма DES.

Аналогичным образом представляется взаимодействие в рассматриваемой игре стратегий A_2 , B_1 и B_2 .

Следуя приведенным выше утверждениям, сформируем платежную матрицу игры [Вентцель Е.С.], представленную в виде Табл. 8.

Таблица 8.

Пользователь «Хакер» \	A_1	A_2
B_1	T_1	$T_1 + T_2$
B_2	$T_2 + T_1$	T_2

Цель игры состоит в том, что Разработчик стремится выбрать стратегию, позволяющую максимизировать затраты времени на «взлом» КП, а «хакер» – стратегию, минимизирующую такие затраты.

Из Табл. 8 видно, что таких единственных (чистых) стратегий у игроков нет, то есть игра не имеет седловой (максиминной) точки для цены игры v [Вентцель Е.С.].

В теории игр доказано, что любая игра с нулевой суммой имеет решение в смешанных стратегиях, когда находящиеся в распоряжении игроков стратегии применяются с некоторыми вероятностями.

Будем описывать такие стратегии для Разработчика с помощью вероятностей p_1 и p_2 применения им стратегий A_1 и A_2 .

Для определения значений p_1 и p_2 и неизвестного значения цены игры v будем решать следующую задачу линейного программирования, сформированную на основе платежной матрицы игры (см. Табл. 8):

$$\begin{aligned} v &\rightarrow \max; \\ T_1 p_1 + (T_1 + T_2) p_2 &\leq v; \\ (T_2 + T_1) p_1 + T_2 p_2 &\leq v; \\ p_1 + p_2 &= 1; \\ 0 \leq p_1 \leq 1, \quad 0 \leq p_2 \leq 1, \quad v &\geq 0. \end{aligned} \tag{163}$$

В связи с малой размерностью сформированной математической модели построим аналитическое решение этой задачи. Это означает, что модель (163) относится к классу аналитически-разрешимых моделей (см. Рис. 13).

Поскольку максимум целевой функции любой задачи линейного программирования достигается на границе множества допустимых решений, перепишем второе и третье ограничения решаемой задачи в виде равенств:

$$\begin{aligned} T_1 p_1 + (T_1 + T_2) p_2 &= v; \\ (T_2 + T_1) p_1 + T_2 p_2 &= v. \end{aligned}$$

Приравнивая их левые части, имеем:

$$T_1 p_1 + (T_1 + T_2) p_2 = (T_2 + T_1) p_1 + T_2 p_2.$$

Преобразуем это выражение следующим образом:

$$p_1(T_1 - T_2 - T_1) = p_2(T_2 - T_1 - T_2).$$

Проводя несложные преобразования, получим равенство вида:

$$p_1 T_2 = p_2 T_1. \quad (164)$$

Искомые переменные p_1 и p_2 должны дополнительно удовлетворять третьему и четвертому ограничениям задачи (163).

Преобразуем третье равенство к виду:

$$p_2 = 1 - p_1$$

и с учетом этого перепишем выражение (164) в следующей форме:

$$p_1 T_2 = (1 - p_1) T_1.$$

Разрешая это уравнение относительно переменной p_1 , получим первое решение задачи вида:

$$p_1^0 = \frac{T_1}{T_1 + T_2}. \quad (165)$$

Второе решение задачи определяется как:

$$p_2^0 = 1 - \frac{T_1}{T_1 + T_2} = \frac{T_2}{T_1 + T_2}. \quad (166)$$

Полученные оптимальные значения вероятностей практически означают, что при эксплуатации КП в $p_1^0 \cdot 100\%$ случаев должен использоваться алгоритм DES и в $p_2^0 \cdot 100\%$ – алгоритм RSA.

Оптимальное значение цены игры v^0 может быть получено путем подстановки решений p_1^0 и p_2^0 в любое из приведенных выше равенств.

Используя первое из них, получим:

$$\begin{aligned} v^0 &= T_1 \frac{T_1}{T_1 + T_2} + (T_1 + T_2) \frac{T_2}{T_1 + T_2} = \\ &= \frac{T_1^2}{T_1 + T_2} + T_2 = \frac{T_1^2 + T_1 T_2 + T_2^2}{T_1 + T_2}. \end{aligned} \quad (167)$$

Отметим, что величина v^0 определяет математическое ожидание (среднее значение) затрат времени «хакера» на «взлом» КП, в которой с вероятностями p_1^0 и p_2^0 применяются алгоритмы DES и RSA.

Рассмотрим вопросы реализации полученных решений в виде физической смеси стратегий A_1 и A_2 .

Потребуем, чтобы минимальное значение длины интервала времени действия алгоритмов DES и RSA определялось как:

$$\Delta t < \min(T_1, T_2).$$

Выберем интервал времени $[0, \tau]$, на котором с вероятностями p_1^0 и p_2^0 должна осуществляться смена алгоритмов DES и RSA. Потенциальное число таких переключений вычисляется по формуле вида:

$$N = \frac{\tau}{\Delta t}.$$

Сформируем на интервале $[0, \tau]$ сетку моментов времени с шагом, равным Δt :

$$0 < t_1 < t_2 < \dots < t_{N-1} < t_N,$$

где узлы сетки вычисляются с помощью следующего рекуррентного выражения:

$$t_j = t_{j-1} + \Delta t, \quad j = \overline{(1, N)}, \quad t_0 = 0.$$

Используя метод индукции, имеем, что:

$$t_j = j \Delta t, \quad j = \overline{(1, N)}.$$

Количество n_1 и n_2 применений на интервале времени $[0, \tau]$ рассматриваемых алгоритмов можно определить из следующих выражений:

$$n_1 = [Np_1^0]; \quad n_2 = [Np_2^0],$$

где $[(\cdot)]$ – операция округления числа (\cdot) до целого значения по «правилу 0,5».

Отметим, что полученные значения должны удовлетворять равенству:

$$n_1 + n_2 = N.$$

Периодичность применения алгоритмов DES и RSA на интервале времени $[0, \tau]$ вычислим следующим образом:

$$\Delta\theta_1 = \frac{\tau}{n_1}; \quad \Delta\theta_2 = \frac{\tau}{n_2}.$$

Определим максимальное значение из этих величин:

$$\Delta\vartheta = \max(\Delta\theta_1, \Delta\theta_2). \quad (168)$$

Сформируем график применения алгоритмов DES и RSA на интервале времени $[0, \tau]$.

При $\Delta\vartheta = \Delta\theta_1$, которая является кратной величине Δt , алгоритм DES включается в моменты времени, определяемые из рекуррентного выражения вида:

$$t_j^{(1)} = t_{j-1}^{(1)} + \Delta\theta_1, \quad j = (\overline{1, n_1}), \quad \text{где } t_0^{(1)} = 0. \quad (169)$$

Во все оставшиеся моменты времени сетки, которые определяются как:

$$t_0^{(2)} = 0, \quad t_j^{(2)} = t_{j-1}^{(2)} + \Delta t, \quad t_j^{(2)} \neq t_j^{(1)}, \quad j = (\overline{1, N})$$

включается алгоритм RSA.

Формулы, аналогичные этим выражениям, можно записать для случая, когда $\Delta\vartheta = \Delta\theta_2$.

При значении $\Delta\vartheta = \Delta\theta_1 = \Delta\theta_2$, кратном Δt , выражение (168) можно применить к любому из алгоритмов DES или RSA.

Если величина $\Delta\vartheta$ не является кратной шагу сетки Δt , то проводится корректировка периодичности применения соответствующего алгоритма. Для этой цели на сетке рассматриваются интервалы времени $[t_{j-1}, t_j]$, $j = (\overline{1, N})$ и из их состава выделяется интервал $[t_{k-1}, t_k]$, для которого выполняется неравенство:

$$t_{k-1} < \Delta\vartheta < t_k, \quad k = \overline{(1, N)}.$$

В этом случае в выражении (169) полагаем $\Delta\vartheta = t_{k-1}$.

В защищаемом КП приведенное выше решающее правило применяется на каждом из интервалов времени его функционирования:

$$[0, \tau], [\tau, 2\tau], [2\tau, 3\tau], \dots \quad (170)$$

при начальных условиях вида:

$$t_0^{(1)} = t_0^{(2)} = 0; \quad t_0^{(1)} = t_0^{(2)} = \tau; \quad t_0^{(1)} = t_0^{(2)} = 2\tau; \dots$$

Рассмотрим примеры, иллюстрирующие применение предлагаемого метода.

Пример.

Пусть на основе сайтов, посвященных вопросам стойкости алгоритмов DES и RSA, получены следующие оценки.

«Взлом» алгоритма DES с помощью компьютерной сети может быть осуществлен за 22 часа. Теоретически «взлом» этого алгоритма возможен за 6 мин на ЭВМ стоимостью 10 млн. дол.

«Взлом» алгоритма RSA при использовании 40-битного кода защиты с помощью офисного компьютера был осуществлен за 13 дней. При использовании процессора P5/100 с интерпретатором be и ОС FreeBSD затраты времени на «взлом» этого алгоритма были равны 20-25 мин. Из приведенных данных оценок выберем их минимальные значения:

$$T_1 = 6 \text{ мин}; \quad T_2 = 20 \text{ мин}.$$

Тогда, используя формулы (165)-(167), имеем:

$$p_1^0 = \frac{6}{6+20} = 0,231; \quad p_2^0 = \frac{20}{6+20} = 0,769;$$

$$v^0 = \frac{6^2 + 6 \cdot 20 + 20^2}{6+20} = 21,38 \text{ мин}.$$

Выигрыш от применения предлагаемого подхода будет равен:

$$\Delta T = \frac{6 + 20}{\max(6, 20)} = 1,3 \text{ раза.}$$

Выберем величину Δt следующим образом:

$$\Delta t = 5 \text{ мин.} < \min(6, 20)$$

Это означает, что в течение 5 минут действующий алгоритм DES не может быть «взломан», а по истечении этого времени осуществляется переключение на алгоритм RSA.

В качестве интервала времени $[0, \tau]$, на котором должна производиться смена действующих алгоритмов DES и RSA, выберем интервал $[0, 60]$ мин. с шагом времени $\Delta t = 5$ мин.

Число таких переключений будет равно:

$$N = \frac{60}{5} = 12.$$

Используя соответствующие выражения и вычисленные выше значения вероятностей p_1^0 и p_2^0 , имеем:

$$n_1 = [12 \cdot 0,231] = [2,772] = 3; \quad n_2 = [12 \cdot 0,772] = [9,264] = 9.$$

Отметим, что для найденных значений условие $n_1 + n_2 = N$ выполняется.

Периодичности применения в течение одного часа алгоритмов DES и RSA будут равны:

$$\Delta\theta_1 = \frac{60}{3} = 20 \text{ мин}; \quad \Delta\theta_2 = \frac{60}{9} = 6,666 \text{ мин.}$$

Максимум $\Delta\theta$ из этих величин соответствует значению $\Delta\theta_1 = 20$ мин, которое является кратным значению $\Delta t = 5$ мин.

Применяя соотношение (169), получаем моменты времени $t_j^{(1)}$, $j = (\overline{1, 3})$, в которые начинает действовать алгоритм DES:

$$t_1^{(1)} = 0 + 20 = 20 \text{ мин}; \quad t_2^{(1)} = 20 + 20 = 40 \text{ мин};$$

$$t_3^{(1)} = 40 + 20 = 60 \text{ мин};$$

Оставшиеся моменты времени $t_j^{(2)}, j = (\overline{0, 9})$, в которые применяется алгоритм RSA конкретизируются как:

$$t_0^{(2)} = 0 \text{ мин}; \quad t_1^{(2)} = 0 + 5 = 5 \text{ мин}; \quad t_2^{(2)} = 5 + 5 = 10 \text{ мин};$$

$$t_3^{(2)} = 10 + 5 = 15 \text{ мин}; \quad t_4^{(2)} = 20 + 5 = 25 \text{ мин};$$

$$t_5^{(2)} = 25 + 5 = 30 \text{ мин}; \quad t_6^{(2)} = 30 + 5 = 35 \text{ мин};$$

$$t_7^{(2)} = 40 + 5 = 45 \text{ мин}; \quad t_8^{(2)} = 45 + 5 = 50 \text{ мин};$$

$$t_9^{(2)} = 50 + 5 = 55 \text{ мин}.$$

Здесь учтено, что на интервалах времени $[15, 20]$ мин., $[35, 40]$ мин., $[55, 60]$ мин. действует алгоритм DES.

График включения в работу модулей КП, реализующих алгоритмы DES и RSA, приведен в Табл. 9.

Таблица 9

t , мин	0	5	10	15	20	25	30	35	40	45	50	55
Алгоритм	RSA	RSA	RSA	DES	RSA	RSA	RSA	DES	RSA	RSA	RSA	DES

Из этой таблицы следует, что время применения в системе алгоритма RSA всегда меньше величины $T_2 = 20$ мин. его «взлома».

Сформируем оценку \bar{v} величины среднего выигрыша Разработчика при использовании предлагаемого подхода:

$$\bar{v} = \frac{n_1}{N} T_1 + \frac{n_2}{N} (T_1 + T_2) = \frac{3}{12} \cdot 6 + \frac{9}{12} \cdot 26 = 21 \text{ мин}.$$

Эта величина отличается от значения v^0 на 0,38 мин., что говорит о практическом их совпадении.

Пример.

Известно, что для повышения стойкости данных осуществляется тройное применение алгоритма DES («3DES»).

Будем считать, что в этом случае величина $T_1 = 3 \times 6 = 18$ мин., $T_2 = 20$ мин.

Применяя для интервала времени $[0, \tau] = [0, 60]$ мин. приведенные выше формулы, получаем следующие результаты:

$$p_1^0 = \frac{18}{18+20} = 0,474; \quad p_2^0 = \frac{20}{18+20} = 0,526;$$

$$v = \frac{18^2 + 18 \cdot 20 + 20^2}{18 + 20} = 28,53 \text{ мин.};$$

$$\Delta T = \frac{18+20}{\max(18, 20)} = 1,9 \text{ раза};$$

$$\Delta t = 15 < \min(18, 20);$$

$$N = \frac{60}{15} = 4;$$

$$n_1 = [4 \cdot 0,474] = [1,88] = 2; \quad n_2 = [4 \cdot 0,526] = [2,12] = 2;$$

$$\Delta\theta_1 = \Delta\theta_2 = \frac{60}{2} = 30 \text{ мин.} = \Delta\theta.$$

Отметим, что полученное $\Delta\theta$ является кратным значению шага сетки $\Delta t = 15$ мин.

График включения в течение одного часа модулей алгоритмов 3DES и RSA представлен в Табл. 10.

Таблица 10

t , мин.	0	15	30	45
Алгоритм	3DES	RSA	3DES	RSA

Из этой таблицы видно, что интервалы времени действия алгоритмов 3DES и RSA меньше величин $T_1 = 18$ мин. и $T_2 = 20$ мин. их «взлома».

Пусть выбрана величина шага сетки $\Delta t = 12$ мин.

В этом случае $N = 5$ и узлы сетки будут иметь следующие значения:

$$t_1 = 12 \text{ мин.}; \quad t_2 = 24 \text{ мин.}; \quad t_3 = 36 \text{ мин.}; \\ t_4 = 48 \text{ мин.}; \quad t_5 = 60 \text{ мин.}$$

Значения n_1 и n_2 будут равны:

$$n_1 = [5 \cdot 0,474] = [2,37] = 2; \quad n_2 = [5 \cdot 0,526] = [2,63] = 3.$$

Периодичности применения алгоритмов 3DES и RSA вычисляются как:

$$\Delta\theta_1 = \frac{60}{2} = 30 \text{ мин.} = \Delta\theta; \quad \Delta\theta_2 = \frac{60}{3} = 20 \text{ мин.}$$

Величина $\Delta\theta$ не является кратной величине шага сетки Δt , то есть частное от деления первой на вторую не является целым числом.

Найдем на сетке значения времени t_j , $j = (\overline{1, 5})$ интервал, содержащий величину $\Delta\theta = 30$ мин. Границы такого интервала согласно приведенному выше условию определяются значениями $t_2 = 24$ мин. и $t_3 = 36$ мин. Тогда при формировании моментов времени применения алгоритмов 3DES величина $\Delta\theta$ полагается равной 24 мин. График применения рассматриваемых алгоритмов представлен в Табл. 11.

Таблица 11

t , мин.	0	12	24	36	48
Алгоритм	3DES	RSA	3DES	RSA	3DES

Из этой таблицы также следует, что интервалы действия алгоритмов не превышают времени их «взлома».

На наш взгляд, приведенный выше детерминированный подход с формированием физической смеси стратегий имеет в качестве недостатка повторяемость процесса переключений алгоритмов DES и RSA на каждом из интервалов (170).

Рассмотрим применение метода Монте-Карло, описанного в работе [Вентцель Е.С.], для реализации оптимальных смешанных стратегий (165), (166). Этот метод позволяет сфор-

мировать реализации случайных событий, заданных значениями вероятностей их наступления в каждом опыте.

В качестве таких событий будем рассматривать случайные в процессе эксплуатации КП события:

A_1^0 – применение алгоритма DES;

A_2^0 – применение алгоритма RSA.

Суть метода Монте-Карло состоит в генерации равномерно распределенного в интервале $(0, 1)$ случайного числа ξ и проверке условия вида:

$$0 < \xi \leq p_1^0. \quad (171)$$

При выполнении этого условия считается, что наступило событие A_1^0 .

Если число ξ таково, что выполняется неравенство:

$$p_1^0 < \xi \leq 1, \quad (172)$$

то предполагается наступление события A_2^0 .

Для практической реализации этого подхода будем использовать бесконечную временную сетку функционирования КП с выбранным выше шагом Δt .

Узлы этой сетки $t_1, t_2, t_3, \dots, t_{j-1}, t_j, \dots$ при $t_0 = 0$ вычисляются по приведенным выше формулам.

На введенной сетке выделим интервалы времени (t_{j-1}, t_j) , $j = 1, 2, 3, \dots$, на которых должны в соответствии с вероятностями p_1^0 и p_2^0 использоваться алгоритмы DES или RSA.

Пусть в момент времени t_{j-1} с помощью датчика случайных чисел (ДСЧ) получено число $\xi_{j-1} \in (0, 1)$, $j = 1, 2, 3, \dots$

Значения ξ_{j-1} и p_1^0 используются при проверке условий (171) и (172). Если выполняется условие (171), то на интервале времени (t_{j-1}, t_j) должен функционировать алгоритм DES, в

противном случае (выполняется условие (172)) на этом интервале времени должен функционировать алгоритм RSA. При этом суммарное время непрерывной работы каждого алгоритма не должно превышать времени его «взлома». Для этого вычисляется максимальное количество k_i следующих подряд интервалов времени Δt , в которых возможно применение одного и того же алгоритма, невзирая на условия (171) и (172):

$$k_i = \begin{cases} \text{entiar}(T_i / \Delta t), & \text{если } (T_i / \Delta t) - \text{entiar}(T_i / \Delta t) > 0; \\ \text{entiar}(T_i / \Delta t) - 1, & \text{если } (T_i / \Delta t) - \text{entiar}(T_i / \Delta t) = 0, \end{cases} \quad i = 1, 2.$$

Здесь функция $\text{entiar}(x)$ выполняет выделение целой части числа x .

Для фиксирования следующих подряд интервалов времени Δt , в которых осуществлялось применение одного и того же алгоритма, используется счетчик таких интервалов m_i , $i = 1$ или 2 .

В каждый момент времени t_{j-1} проверяется условие:

$$m_i \leq k_i, \quad i = 1 \text{ или } 2.$$

Если это неравенство выполняется, то на интервале (t_{j-1}, t_j) применяется ранее используемый алгоритм и к счетчику m_i прибавляется единица. В противном случае в момент времени t_{j-1} происходит смена алгоритма, и соответствующий счетчик m_i обнуляется.

Результаты трех вычислительных экспериментов с исходными данными:

$$p_1^0 = 0,474; \quad p_2^0 = 0,526; \quad T_1 = 18; \quad T_2 = 20;$$

$$\Delta t = 15; \quad k_1 = 1; \quad k_2 = 1$$

представлены в Табл. 12.

Таблица 12

Интервалы времени	Вариант 1		Вариант 2		Вариант 3	
	Случ. число	Алгоритм	Случ. число	Алгоритм	Случ. число	Алгоритм
0-15	0,5296	RSA	0,517	RSA	0,2085	DES
15-30	0,2771	DES	0,3579	DES	0,1637	RSA
30-45	0,8499	RSA	0,7126	RSA	0,6433	DES
45-60	0,0726	DES	0,5946	DES	0,9608	RSA
60-75	0,6542	RSA	0,4412	RSA	0,1768	DES
75-90	0,3204	DES	0,2781	DES	0,5347	RSA
90-105	0,9154	RSA	0,11	RSA	0,7989	DES
105-120	0,7891	DES	0,5332	DES	0,8646	RSA

Самостоятельная работа.

Разработать математическую модель, аналогичную выражению (163), для случая применения в ИС n методов шифрования и дешифрования применяемых данных.

Решить рассмотренную задачу с применением случайного (Лекция №8) или стохастического (Лекция №11) бинарного отношения.

В заключение лекции отметим, что достаточно глубоко разработанные к настоящему времени модели теории игр, рассматривающей принятие решений в ситуациях, когда ЛПР имеют противоположные цели, могут стать хорошей основой для создания перспективных методов обеспечения эффективной защиты информационных систем.

23. Дискретная модель активной защиты информационных систем

Как показал анализ современного состояния проблемы информационной безопасности (ИБ), общепринятые в настоящее время подходы к защите информационных систем

(ИС) ставят основной целью нейтрализацию и минимизацию их потерь от информационных воздействий(атак) противника. Это позволяет считать их *пассивными методами и средствами обеспечения ИБ* таких систем.

На наш взгляд, дополнительного повышения эффективности ИБ можно добиться путём применения активных методов и средств защиты ИС, осуществляющих полную или частичную нейтрализацию выявленных источников информационных атак.

Следуя вероятностному подходу к оценке безопасности ИС, покажем эффективность таких методов и средств.

В качестве показателей эффективности защиты системы A при воздействиях(атаках) на неё некоторой системы B будем рассматривать вероятности нарушения ее ИБ при использовании системой A пассивных (традиционных) $P_{\text{ПИБ}}^A$ и активных $P_{\text{АИБ}}^A$ методов и средств защиты.

Пусть D - случайное событие, определяющие нарушение ИБ системы A средствами нападения (СН) системы B . Состояние дееспособности этих средств будем описывать с помощью случайного события E .

При этом предполагается, что вероятность наступления события E может принимать следующие значения: $P(E) = 1$, если СН полностью дееспособны; $P(E) < 1$, если СН частично дееспособны; $P(E) = 0$, если СН полностью не дееспособны, то есть выведены из строя средствами защиты стороны A .

Используя формулу произведения вероятностей [Гнеденко Б.В.] имеем, что

$$P_{\text{АИБ}}^A = P(D | E)P(E).$$

Пассивные методы и средства защиты обычно применяются самой системой A в предположении, что СН стороны B полностью дееспособны, то есть $P(E) = 1$. В этом случае получаем, что:

$$P_{\text{ПИБ}}^A = P(D).$$

Реальным условиям информационного противоборства соответствуют неравенства вида:

$$0 < P(E) < 1, \quad P(D|E) \leq P(D).$$

Отсюда следует, что:

$$P_{\text{АИБ}}^A < P_{\text{ПИБ}}^A,$$

то есть методы и средства активной защиты, применяемые системой A , являются более эффективными по сравнению с пассивными методами и средствами. На это указывает и тот факт, что при $P(E) = 0$ величина $P_{\text{АИБ}}^A = 0$, то есть нарушение ИБ системы A не происходит.

Рассмотрим основные вопросы организации средств и методов активной защиты ИС. Как известно, одной из важнейших задач информационного противоборства является вывод из строя вычислительных мощностей противника [Моисеев В.С., Козар А.Н., Дятчин В.В.]. Другой целью информационных атак может быть внесение дезинформации в базы данных противника [Герасименко А.В.].

Будем считать, что в перспективных ИС задачи активной защиты должны решаться специальными аппаратно-программными средствами, которые будем называть *серверами активной защиты (САЗ)*. Такие серверы наряду с функциями пассивной защиты, такими как:

- организация и поддержание схем доступа в соответствии с уровнями ИБ,

- реализация криптографических методов обеспечения ИБ,
- антивирусная защита ИС,
- межсетевое экранирование,
- физическая защита ИС,
- скрывание информации на основе компьютерной стенографии, должны решать следующие дополнительные задачи:

- 1) идентификация атакующей системы и типов атак с применением средств компьютерной разведки [Жарова О.Ю., Федорова В.А.],

- 2) выбор видов и средств реализации информационных атак,

- 3) вскрытие средств защиты противника,

- 4) мониторинг поведения захваченных в результате ответных информационных атак средств защиты противника.

Общая схема информационного противоборства систем A и B , использующих средства активной защиты своих информационных, программных и аппаратных ресурсов, приведена на Рис. 73.

Для дальнейшего рассмотрения будем считать, что ИС сторон A и B оснащены множествами САЗ, включающими в себя n и m серверов соответственно. Предполагается, что эти стороны последовательно осуществляют серию «дуэльных» информационных атак до полного вывода из строя САЗ одной из сторон.

Отметим, что при высоких требованиях к ИБ защищаемых ИС, значительных объёмах обрабатываемой в них информации и жестких временных ограничениях на решение, приведённых выше задач активной защиты их ресурсов, в составе САЗ предлагается использовать такие специализированные

аппаратно-программные средства как разведывательные (РЗС) и командные серверы (КС).

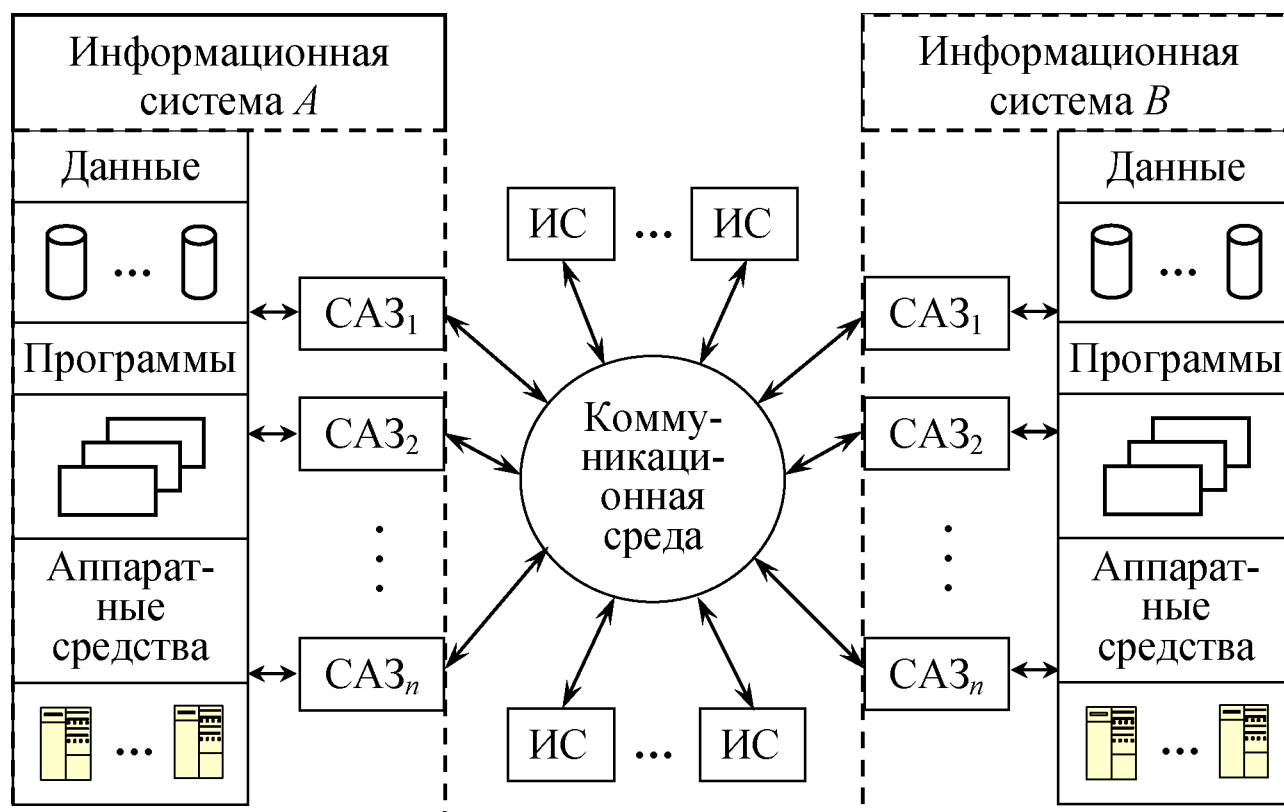


Рис. 73

Основным назначением РЗС является реализация задач поиска серверов противника, установление их адресов и типов атак, отслеживание результативности проводимых на них атак и мониторинг их активности в случае успешной реализации скрытого захвата управления (СЗУ) над ними.

Специальные КС, действующие по данным РЗС предназначены для непосредственного проведения информационных атак по уничтожению (захвату) средств противника, для нарушения доступности и целостности информационных ресурсов противника, а так же для размещения на них дезинформации и скрытых программных модулей (проведения операций СЗУ).

Отметим, что для создания эффективных методов и средств активной защиты ИС необходима разработка её теоретических основ, в частности математических моделей анализа и оптимального синтеза таких средств.

В данной лекции рассмотрим одну из таких моделей, описывающую процесс противоборства САЗ сторон A и B , которая основана на использовании формализма «задачи о разорении игроков» [Гнеденко Б.В.].

Обозначим через p и q вероятности успешных атак стороны A и стороны B . Потребуем, чтобы их значения удовлетворяли условиям:

$$p + q = 1, \quad p \in (0,1), \quad q \in (0,1), \quad (173)$$

подразумевающим отсутствие ничьих в каждой из атак.

Предполагается, что результатом каждой атак, которые считаются независимыми, является захват той или иной стороной одного САЗ противника, который может быть использован для нанесения его ИС определённого ущерба.

Одной из задач анализа эффективного применения методов и средств активной защиты является определение вероятностей поражения сторон A и B при заданных значениях параметров n , m , p и q .

Обозначим через P_k и Q_r вероятности поражения сторон A и B при наличии у них соответственно k и r действующих в их интересах САЗ.

Отсюда вероятности поражения этих сторон в серии «дуэльных» информационных атак определяются значениями P_n и Q_m .

Если сторона A перед некоторой атакой имела k серверов, то её полное поражение в информационной борьбе со стороной B может осуществиться двумя путями: очередная атака

стороны A будет успешной, но все остальные она проиграет, или она потерпит поражение в текущей атаке и во всем противоборстве. Исходя из этого по формуле полной вероятности [Гнеденко Б.В.] получим:

$$P_k = pP_{k+1} + qP_{k-1}.$$

Аналогичное выражение для стороны B записывается как

$$Q_r = qQ_{r+1} + pQ_{r-1}.$$

Представим эти выражения в виде линейных однородных конечно-разностных уравнений 2-го порядка с постоянными коэффициентами [Миронов А.А., Солдатов М.А.]:

$$\begin{aligned} pP_{k+1} - P_k + qP_{k-1} &= 0; \\ qQ_{r+1} - Q_r + pQ_{r-1} &= 0. \end{aligned} \tag{174}$$

Следуя этой работе общие решения полученных уравнений при $p \neq q$ записываются в следующем виде:

$$\begin{aligned} P_k &= C_1 \lambda_1^k + C_2 \lambda_2^k; \\ Q_r &= C_3 \mu_1^r + C_4 \mu_2^r. \end{aligned}$$

В этих выражениях C_i , $i = (\overline{1, 4})$ это постоянные, определяемые из начальных или граничных условий уравнений (174), а λ_1, λ_2 и μ_1, μ_2 - корни характеристических уравнений вида:

$$p\lambda^2 - \lambda + q = 0; \quad q\mu^2 - \mu + p = 0.$$

Решая эти уравнения с использованием условия (173), получим, что:

$$\lambda_1 = 1; \quad \lambda_2 = \frac{q}{p}; \quad \mu_1 = \frac{p}{q}; \quad \mu_2 = 1.$$

Тогда общие решения уравнений (174) конкретизируются следующим образом:

$$P_k = C_1 + C_2 (q/p)^k, \quad Q_r = C_3 + C_4 (p/q)^r. \tag{175}$$

Для $p = q$ общие решения уравнений (174) будут иметь вид [Миронов А.А., Солдатов М.А.]:

$$P_k = C_5 + C_6 k, \quad Q_r = C_7 + C_8 r. \quad (176)$$

Для определения значений постоянных C_i , $i = (\overline{1, 8})$, входящих в выражения (175), (176) сформируем граничные условия для уравнений (174).

Будем считать, что случайными событиями, описывающими абсолютную победу стороны A или B , является захват всех серверов противника при сохранении всех своих серверов. Тогда в случае поражения стороны A имеем, что:

$$P_{n+m} = 0. \quad (177)$$

Кроме того, поражение стороны A объективно происходит, когда она не имеет ни одного действующего («незахваченного» стороной B) сервера. Это условие записывается как:

$$P_0 = 1. \quad (178)$$

Из аналогичных рассуждений следует вид граничных условий для второго из уравнений (174):

$$Q_{n+m} = 0; \quad Q_0 = 1. \quad (179)$$

Используя граничные условия (177) и (178) применительно к первому из выражений (175) и определяя значения C_1 и C_2 из решения полученной системы уравнений, частное решение первого из уравнений (174) можно представить выражением вида:

$$P_k = \frac{(q/p)^{n+m} - (q/p)^k}{(q/p)^{n+m} - 1}, \quad k = (\overline{0, n+m}). \quad (180)$$

Частное решение второго уравнения, полученное аналогичным образом с использованием второго из выражений (175) и формулы (179) записывается как:

$$Q_r = \frac{(p/q)^{n+m} - (p/q)^r}{(p/q)^{n+m} - 1}, \quad r = (\overline{0, n+m}). \quad (181)$$

Частные решения уравнений (174) при $p = q = 0,5$, построенные на основе выражений (176) и граничных условий (177)-(179) имеют вид:

$$P_k = 1 - \frac{k}{n+m}; \quad Q_r = 1 - \frac{r}{n+m}, \quad (182)$$

$$k = (\overline{0, n+m}), \quad r = (\overline{0, n+m}).$$

Искомые вероятности P_n и Q_m поражения сторон A и B при $p \neq q$, получаемые при подстановке в формулы (180), (181) значений $k = n$ и $r = m$ и проведения соответствующих преобразований, могут быть определены с помощью следующих формул:

$$P_n = \frac{1 - (p/q)^m}{1 - (p/q)^{n+m}}; \quad Q_m = \frac{1 - (q/p)^n}{1 - (q/p)^{n+m}}. \quad (183)$$

Для случая $p = q = 0,5$ эти вероятности с использованием формул (173) вычисляются как:

$$P_n = \frac{m}{n+m}; \quad Q_m = \frac{n}{n+m}. \quad (184)$$

Пример.

Пусть у сторон A и B имеется по два САЗ (см. Рис. 73), то есть $n = 2$, $m = 2$. При этом известно, что вероятности их успешных атак имеют следующие значения: $p = 0,8$ и $q = 0,2$. С учётом этих исходных данных, применяя формулы (183), получаем следующие значения вероятностей поражения сторон A и B в их информационном противоборстве:

$$P_2 = 0,0588; \quad Q_2 = 0,9412.$$

Эти результаты показывают, что за счёт того, что $p \gg q$ сторона B с вероятностью близкой к единице потерпит поражение при равном со стороной A числе САЗ.

Если $p = q = 0,5$, то согласно формулам (184) вероятности $P_2 = Q_2 = 0,5$. Из этих формул можно сделать вывод о том, что при $n = m$ вероятности поражения сторон A и B будут всегда равняться 0,5.

Для значений $n = 1$, $m = 2$, $p = 0,85$ и $q = 0,15$ с использованием выражений (183) соответствующие вероятности определяются как:

$$P_1 = 0,1719, \quad Q_1 = 0,8281.$$

Из этих результатов следует, что сторона B с большей вероятностью потерпит поражение, чем сторона A за счет того, что $p > q$, хотя $n < m$.

Проведём исследование полученных зависимостей и рассмотрим вопросы их практического применения при реализации активной защиты ИС. Из выражений (180)-(182) следует, что всегда выполняется условие вида:

$$P_n + Q_m = 1. \quad (185)$$

Вероятности «побед» сторон A и B определяются как:

$$P_n^* = 1 - P_n, \quad Q_m^* = 1 - Q_m.$$

Тогда с учётом условия (185) получаем, что $P_n^* = Q_m$, $Q_m^* = P_n$.

Построим оценку вероятности успешных атак стороны A , оснащённой n САЗ, обеспечивающей поражение стороны B .

Условием поражения стороны B является выполнение неравенства $Q_m > P_n$. Используя приведённое выше условие нормировки этих вероятностей (185) получаем, что:

$$Q_m > 0,5. \quad (186)$$

При $p = q = 0,5$ с подстановкой в это неравенство второго выражения из формулы (184) условие поражения стороны B конкретизируется как:

$$n > m.$$

Для случая $p \neq q$ подставим в левую часть неравенства (186) выражение (181) при $r = m$:

$$\frac{(p/q)^{n+m} - (p/q)^m}{(p/q)^{n+m} - 1} > \frac{1}{2}.$$

Введём вспомогательную переменную $x = p/q$ и перепишем это неравенство в виде:

$$\frac{x^{n+m} - x^m}{x^{n+m} - 1} > \frac{1}{2}.$$

Выполним следующее преобразование:

$$\frac{1}{2}x^{n+m} - x^m + \frac{1}{2} > 0,$$

Не ослабляя это неравенство, отбросим положительный свободный член и получим неравенство вида:

$$\frac{1}{2}x^{n+m} - x^m > 0.$$

Проведя ряд простых преобразований с учётом условия (173) и замены $x = p/q$ получим, что для поражения стороны B сторона A при наличии у неё n серверов должна осуществлять успешные атаки с вероятностью:

$$p > \frac{2^{1/n}}{1 + 2^{1/n}} = p^*. \quad (187)$$

Известно, что величина $2^{1/n} > 1$ при любом значении $n \geq 1$. Тогда из выражения (187) следует, что сторона A должна обладать вероятностью успешных атак $p > 0,5$.

Определим нижнюю границу числа n серверов стороны A , обеспечивающих при фиксированных значениях вероятностей p и q ($p > q$) нанесение поражения стороне B с ее m серверами, с вероятностью не меньше чем α . Это требование записывается в виде следующего неравенства:

$$Q_m \geq \alpha. \quad (188)$$

Для случая $p = q = 0,5$, подставляя в левую часть этого неравенства второе выражение из формулы (184) после несложных преобразований получим, что для «победы» стороны A с вероятностью не меньшей чем α , она должна иметь число серверов, удовлетворяющих неравенству вида:

$$n \geq \left[\frac{\alpha m}{1 - \alpha} \right], \quad (189)$$

где $[(\cdot)]$ – целая часть числа (\cdot) , округлённого до большего целого.

Пример.

Пусть $m = 4$ и $\alpha = 0,8$. Тогда из неравенства (189) следует, что $n \geq 16$. Если $\alpha = 0,9$, то требуемое число серверов стороны A должно быть $n \geq 36$, а в случае $m = 10$ и $\alpha = 0,9$, их число определяется как $n \geq 90$.

Для случая $p \neq q$ неравенство (188) с учетом выражения для Q_m формулы (183) примет вид:

$$\frac{1 - (q/p)^n}{1 - (q/p)^{n+m}} \geq \alpha.$$

Решая это неравенство относительно величины n , получим искомую оценку вида:

$$n \geq \left[\frac{\ln(1 - \alpha) - \ln(1 - \alpha(q/p)^m)}{\ln(q/p)} \right].$$

Будем считать, что $\alpha = 0,9$ и вероятность p изменяется в пределах от 0,55 до 0,95 с шагом 0,05. Построим, используя эту формулу, нижние значения оценки числа n серверов стороны A для числа серверов противника $m = \{1; 2; 4; 10\}$.

Результаты вычислительных экспериментов представлены на Рис. 74.

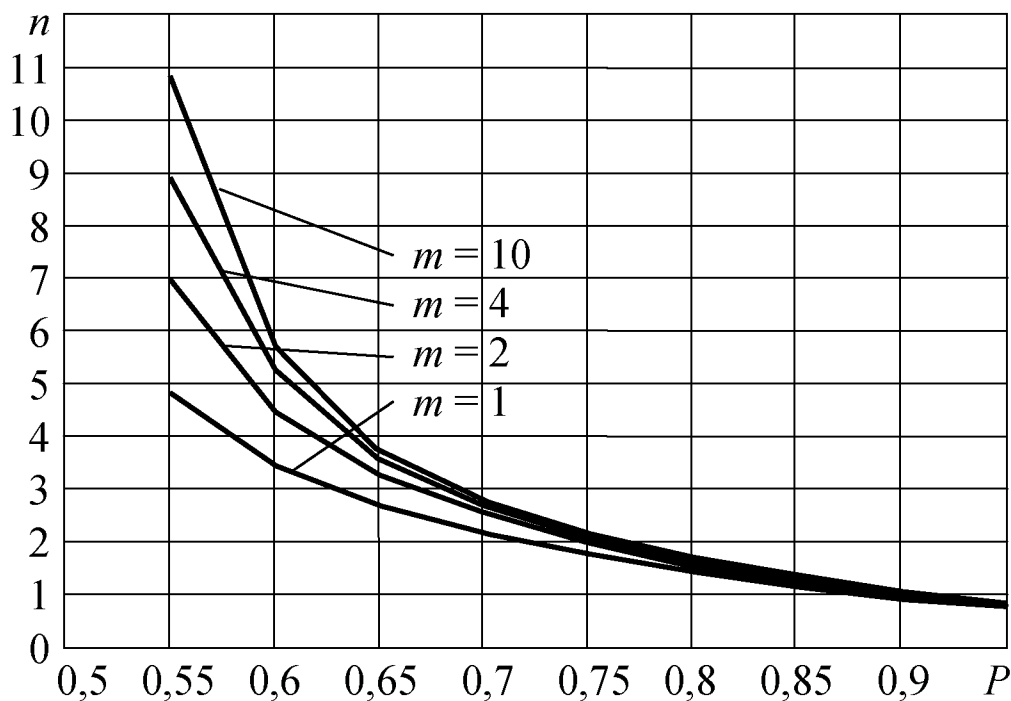


Рис. 74

Следует отметить, что оценка $n \geq 11$ для $p = 0,55$ и $m = 4$ и $m = 10$ является более приемлемой для практики, чем оценки, полученные выше. Также из этого рисунка следует, что вне зависимости от числа m САЗ стороны B при $p \in [0,8; 0,9]$ минимальное потребное число серверов стороны A будет равно $n(p) = 2$, а при $p = 0,95$ — $n(p) = 1$.

Таким образом, средства активной защиты стороны A для «победы» над стороной B с вероятностью $\alpha = 0,9$ должны иметь не менее двух САЗ с вероятностями p успешных атак не менее величины 0,8.

Отметим, что разработанная в этой лекции модель (174) относится к классу аналитически-разрешимых моделей (см. Рис. 13).

Самостоятельная работа.

Разработать имитационную модель информационного противоборства сторон A и B , основанную на применении понятия стохастического отношения (см. Лекцию №11), построенного на множествах W и V их САЗ.

В заключение лекции отметим, что методология активной защиты ИС нуждается в проведении более глубокой разработки.

24. Марковская модель функционирования программных средств активной защиты информационных систем

Целью данной лекции является конкретизация функций средств активной защиты в рамках их применения для обеспечения информационной безопасности (ИБ) информационных систем (ИС) и анализ функционирования ИС в случае применения в них средств активной защиты. При изложении ее материала воспользуемся статьей [Моисеев В.С., Тутубалин П.И.].

Вопросы формирования состава программного обеспечения средств активной защиты являются особенно актуальными для такого перспективного вида автоматизированных систем управления специального назначения [Моисеев В.С., Козар А.Н., Дятчин В.В.], как мобильные распределенные автоматизированные системы управления (МРАСУ). Эти системы являются отдельным практически неотраженным в существующей литературе видом современных ИС. Отметим, что характерными особенностями таких систем являются: актив-

ное использование в процессе работы различных современных инфокоммуникационных технологий и относительно небольшое время функционирования в каждом новом месте их развертывания (от 2-3 мин. до нескольких часов). Последнее подразумевает высокую степень неопределенности действий противника по нарушению функционирования МРАСУ и объективный дефицит времени на их отражение. В этой связи резко возрастает роль компьютерной разведки, оперативно выявляющей действующие угрозы системе и уязвимости в средствах нападения противника.

Отмеченные выше особенности, а также вероятностный характер результатов компьютерной разведки определяют стохастическую природу процессов активной защиты информационных систем и, в частности, МРАСУ.

Введем в рассмотрение следующие функции (задачи), процесса активной защиты информационной системы:

1. Дистанционный мониторинг вторжений противника в работу элементов системы.
2. Анализ характера вторжений и определение адресов (координат) средств нападения на ИС.
3. Поиск уязвимостей действующих против ИС средств противника.
4. Анализ возможностей использования выявленных уязвимостей.
5. Выбор средств и реализация ответных атак.

Для реализации каждой из этих задач целесообразным, с точки зрения безопасности и отказоустойчивости, будет применение специализированных комплексов программ, входящих в состав функционального программного обеспечения средств активной защиты, которые функционируют в среде

серверов активной защиты (САЗ). Каждому комплексу программ (КП) назначим номер проведенных выше задач, которую он решают.

Отметим, что комплексы программ №4 и №5 должны функционировать в интерактивном режиме под управлением администратора безопасности ИС.

Алгоритмы работы комплексов программ №1 и №2 предлагается реализовать путем соответствующих доработок методов и средств, описанных в работе [Лукацкий А.В.].

В основу разработки комплексов программ №3 и №4 могут быть положены модифицированные к применению в составе ИС методы, представленные в существующей литературе, например, [Скембей Дж., Мак-Клар Ст., Курц Дж. и др.].

При реализации комплекса программ №5 предлагается использовать наряду с существующими такие перспективные методы подавления средств противника, как DOS (DDOS)-атаки, компьютерные вирусы, «радиовирусы», постановщики помех, «боннеты» и др.

Отметим, что по классификации работы [Моисеев В.С., Козар А.Н., Дятчин В.В.] указанные выше программно-аппаратные средства активной защиты являются образцами оборонительного информационного оружия.

Будем считать, что каждый i -й комплекс программ имеет случайное время работы, распределенное по показательному закону с математическим ожиданием T_i , $i = \overline{1, 5}$.

Отметим, что показательный закон выбран как наиболее просто реализуемый в условиях начального этапа исследования предлагаемой модели. Взаимодействие компонент программного обеспечения средств активной защиты может быть

представлено графом связи комплексов программ (КП), который приведен на Рис. 75.

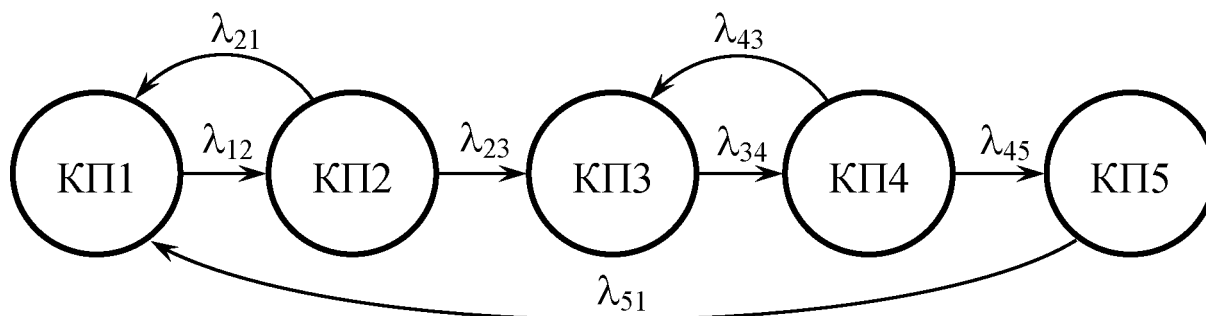


Рис. 75

На этом рисунке параметры λ_{ij} описывают интенсивности передач управления между i -м и j -м комплексами программ, $i, j \in \overline{1, 5}$ в процессе решения задач активной защиты ИС.

Следуя работе [Вентцель Е.С.], параметры λ_{ij} будем вычислять по формулам вида:

$$\begin{aligned} \lambda_{12} &= 1/T_1; \lambda_{21} = (1-\alpha)/T_2; \lambda_{23} = \alpha/T_2; \lambda_{34} = 1/T_3; \\ \lambda_{43} &= (1-\beta)/T_4; \lambda_{45} = \beta/T_4; \lambda_{51} = 1/T_5; \end{aligned} \quad (190)$$

где α – вероятность того, что имеющееся вторжение противника в работу ИС будет идентифицировано средствами активной защиты; β – вероятность того, что выявленная уязвимость в средствах нападения противника может быть использована для организации ответных атак.

Значения параметров, входящих в выражения (190), при создании средств активной защиты могут быть заданы экспертным путем, то есть с использованием группы экспертов в области ИБ систем. Отметим также, что на завершающей стадии разработки ИС они могут определяться по результатам испытаний программного обеспечения средств активной защиты.

Процесс функционирования комплексов программ активной защиты ИС будем описывать марковским процессом с

конечным множеством состояний КП1, КП2,..., КП5 и непрерывным временем [Вентцель Е.С.].

Пусть $p_i(t)$ – вероятность того, что в момент времени t в среде серверов активной защиты ИС выполняется i -й комплекс программ, где $i = (\overline{1,5})$. Динамику изменения на интервале времени $[0, \tau]$ вероятностей $p_i(t)$, учитывая граф связей состояний процесса (см. Рис. 75), будем описывать следующей системой дифференциальных уравнений Колмогорова [Вентцель Е.С.]:

$$\begin{aligned}\dot{p}_1 &= -\lambda_{12}p_1 + \lambda_{21}p_2 + \lambda_{51}p_5, & 0 \leq t \leq \tau; \\ \dot{p}_2 &= \lambda_{12}p_1 - (\lambda_{21} + \lambda_{23})p_2; \\ \dot{p}_3 &= \lambda_{23}p_2 - \lambda_{34}p_3 + \lambda_{43}p_4; \\ \dot{p}_4 &= \lambda_{34}p_3 - (\lambda_{43} + \lambda_{45})p_4; \\ \dot{p}_5 &= \lambda_{45}p_4 - \lambda_{51}p_5,\end{aligned}\tag{191}$$

Отметим, что для МРАСУ τ – время функционирования системы в месте ее дислокации.

Начальные условия для этой системы уравнений имеют вид:

$$p_1(0) = 1, \quad p_i(0) = 0, \quad i = (\overline{2,5}).\tag{192}$$

Вероятности $p_i(t)$ должны удовлетворять следующим условиям:

$$\begin{aligned}0 \leq p_i(t) \leq 1, \quad i = (\overline{1,5}), \quad t \in [0, \tau], \\ \sum_{i=1}^5 p_i(t) = 1, \quad t \in [0, \tau].\end{aligned}\tag{193}$$

Динамическая модель (191)-(193) является переопределенной, т.к. включает в себя пять искомых функций $p_i(t)$ и шесть уравнений для их определения.

Преобразуем выражение (193) к виду:

$$p_5(t) = 1 - \sum_{i=1}^4 p_i(t), \quad (194)$$

подставим его в первое уравнение системы (191) и исключим из рассмотрения ее последнее уравнение.

В этом случае получаем систему уравнений 4-го порядка:

$$\begin{aligned} \dot{p}_1 &= -\lambda_{12}p_1 + \lambda_{21}p_2 + \lambda_{51}\left(1 - \sum_{i=1}^4 p_i\right), \quad t \in [0, \tau]; \\ \dot{p}_2 &= \lambda_{12}p_1 - (\lambda_{21} + \lambda_{23})p_2; \\ \dot{p}_3 &= \lambda_{23}p_2 - \lambda_{34}p_3 + \lambda_{43}p_4; \\ \dot{p}_4 &= \lambda_{34}p_3 - (\lambda_{43} + \lambda_{45})p_4 \end{aligned} \quad (195)$$

с начальными условиями

$$p_1(0) = 1, \quad p_i(0) = 0, \quad i = (\overline{2, 4}). \quad (196)$$

Тогда вычисление искомых вероятностей $p_i(t), i = (\overline{1, 5})$ состоит в решении задачи Коши (195), (196) и применении расчетной формулы (194).

В связи с тем, что марковский процесс, описывающий взаимодействие компонент программного обеспечения средств активной защиты, не имеет поглощающих состояний (см. Рис. 75), все комплексы программ, начиная с некоторого момента времени $t = \tau^* \leq \tau$, могут также функционировать в установившемся режиме с постоянными вероятностями $p_1^*, p_2^*, \dots, p_5^*$ [Вентцель Е.С.].

Предложенная выше модель функционирования программного обеспечения средств активной защиты ИС позволяет определить ряд вероятностно-временных характеристик, необходимых для организации процесса выполнения КП №1, КП №2, ..., КП №5 на серверах активной защиты ИС.

Заметим, что правые части каждого из уравнений линейной системы дифференциальных уравнений (191) включают в

себя коэффициенты с противоположными знаками. Это позволяет предположить, что производные $\dot{p}_i(t)$ при изменении значений функций $p_i(t)$ в пределах от 0 до 1 будут изменять свои знаки и в некоторые моменты времени $t = t_i^* \in [0, \tau]$ могут обращаться в ноль. Последнее означает наличие точек экстремума вероятностей $p_i(t)$, $i = (\overline{1, 5})$. Значения t_i^* определяют моменты времени минимального или максимального использования при активной защите ИС i -го комплекса программ, где $i = (\overline{1, 5})$.

Исходя из марковости применяемой модели, случайные величины $N_i(t)$, описывающие количество передач управления в момент времени $t \in [0, \tau]$ i -му комплексу программ, будут подчинены пуассоновским законам распределения [Вентцель Е.С.], которые описываются вероятностями вида:

$$Q_i(t, k) = P\{N_i(t) = k\} = \frac{(\Lambda_i t)^k}{k!} e^{-\Lambda_i t}, \quad (197)$$

$$i = (\overline{1, 5}), \quad t \in [0, \tau], \quad k = 0, 1, 2, \dots$$

Из Рис. 75 следует, что входящие в эту формулу интенсивности Λ_i вычисляются с учетом выражений (190) как:

$$\Lambda_1 = \lambda_{21} + \lambda_{51}; \quad \Lambda_2 = \lambda_{12};$$

$$\Lambda_3 = \lambda_{23} + \lambda_{43}; \quad \Lambda_4 = \lambda_{34}; \quad \Lambda_5 = \lambda_{45}.$$

Определим математические ожидания $m_i(\tau)$ и дисперсии $\sigma_i^2(\tau)$ общих затрат времени работы на интервале $[0, \tau]$ каждого i -го комплекса программ, где $i = (\overline{1, 5})$.

Учитывая выражения (195), (196) и (194), используемые для определения вида функций $p_i(t)$, запишем формулы вида:

$$m_i(\tau) = \int_0^\tau p_i(t) dt; \quad \sigma_i^2(\tau) = \int_0^\tau p_i(t)(1 - p_i(t)) dt, \quad i = (\overline{1, 5}), \quad (198)$$

Приведем основные свойства математического ожидания $m_i(\tau)$, $i = (\overline{1, 5})$. Из условий того, что $p_i(t) \in [0; 1]$ следует неравенство вида:

$$m_i(\tau) \leq \int_0^{\tau} 1 \cdot dt = \tau.$$

Покажем, что:

$$\sum_{i=1}^5 m_i(\tau) = \tau.$$

Используя расчетную формулу для вычисления $m_i(\tau)$ и условие (193), имеем:

$$\sum_{i=1}^5 m_i(\tau) = \sum_{i=1}^5 \int_0^{\tau} p_i(t) dt = \int_0^{\tau} \left(\sum_{i=1}^5 p_i(t) \right) dt = \int_0^{\tau} 1 \cdot dt = \tau.$$

С помощью выражений (198) и известного «правила 3σ » можно построить интервальные оценки $(m_i(\tau) \pm 3\sigma_i(\tau))$ фактических затрат времени работы i -го комплекса программ на интервале времени $[0, \tau]$ функционирования ИС. При этом с использованием неравенства Чебышева [Гнеденко Б.В.] можно показать, что достоверность таких оценок будет не менее чем 0,889.

Пример.

Пусть средства активной защиты некоторой подсистемы гипотетической МРАСУ специального назначения, функционирующей не более 5 минут, описываются следующими исходными данными:

$$\tau = 300 \text{ с}; T_1 = 12 \text{ с}; T_2 = 8 \text{ с}; T_3 = 20 \text{ с};$$

$$T_4 = 15 \text{ с}; T_5 = 5 \text{ с}; \alpha = 0,8; \beta = 0,7.$$

Результаты решения задачи Коши (195), (196) при значениях интенсивностей, вычисленных по формулам (190) и применении выражения (194), представлены на Рис. 76, из ко-

торого следует, что наиболее используемыми на интервале времени $[0; 360]$ с являются комплексы программ №1 и №3.

Также из полученных результатов следует, что:

а) при $t_1^* = 42,9$ с вероятность $p_1(t)$ достигает минимального значения $p_{1 \min} = 0,171$; точки $t_2^* = 10,5$ с; $t_3^* = 33,6$ с; $t_4^* = 61,5$ с, $t_5^* = 69,3$ с являются точками максимума вероятностей $p_2(t)$, $p_3(t)$, $p_4(t)$, $p_5(t)$ со значениями: $p_{2 \max} = 0,307$, $p_{3 \max} = 0,400$, $p_{4 \max} = 0,273$, $p_{5 \max} = 0,064$.

б) начиная с момента времени $\tau^* \approx 90$ с программное обеспечение средств активной защиты функционирует в установившемся режиме с вероятностями $p_1^* = 0,178$; $p_2^* = 0,125$; $p_3^* = 0,357$; $p_4^* = 0,268$; $p_5^* = 0,063$. Это подтверждает предположения о наличии экстремальных и установившихся значений вероятностей $p_i(t)$, $i = \overline{1,5}$.

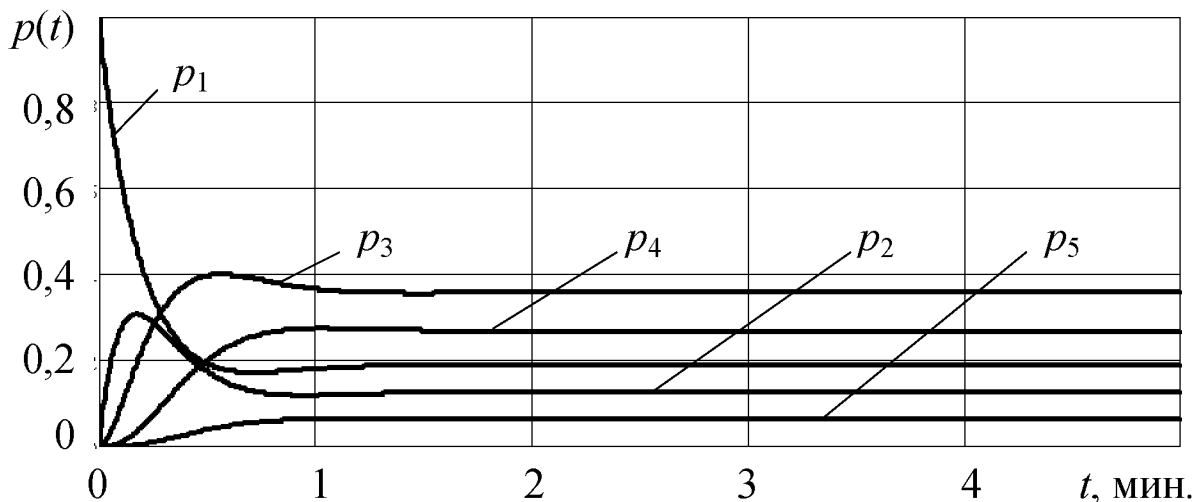


Рис. 76

Результаты анализа влияния значений вероятностей $\alpha \in [0,5; 1]$ и $\beta \in [0,5; 1]$ на вероятность $p_5(300)$ представлены в Табл. 13.

Таблица 13

$\alpha \backslash \beta$	0,5	0,6	0,7	0,8	0,9	1
0,5	0,043	0,046	0,048	0,050	0,051	0,053
0,6	0,048	0,052	0,054	0,057	0,058	0,06
0,7	0,053	0,057	0,060	0,063	0,065	0,067
0,8	0,056	0,061	0,065	0,068	0,070	0,073
0,9	0,060	0,065	0,069	0,073	0,076	0,078
1	0,063	0,068	0,073	0,077	0,080	0,083

Из этой таблицы следует, что при изменении параметров α и β от их минимальных до максимальных значений вероятности нанесения противнику ответного удара в момент времени $\tau = 300$ с увеличивается практически в 2 раза.

Отсюда можно сделать вывод о необходимости максимизации вероятностей α и β при разработке программного обеспечения средств активной защиты.

При использовании выражений (197) при $i = (1,5)$, $t = 300$ с и $k = 0, 1, 2, \dots, 60$ было установлено, что вероятности $Q_2(t, k), \dots, Q_5(t, k)$ имеют явно выраженные максимальные значения по числу k передач управления соответствующим комплексам программ.

Результаты проведенных вычислительных экспериментов имеют вид: комплекс программ №2: $k_{\max} = 25$; $Q_{2 \max} = 0,0795$; комплекс программ №3: $k_{\max} = 35$; $Q_{2 \max} = 0,0663$; комплекс программ №4: $k_{\max} = 15$; $Q_{2 \max} = 0,1024$; комплекс программ №5: $k_{\max} = 14$; $Q_{2 \max} = 0,1059$.

Вероятность $Q_1(300, k)$ с ростом k возрастает от 0 до 0,0334 и достигает наибольшего значения при $k = 60$.

Таким образом, наиболее используемыми при функционировании средств активной защиты рассматриваемой подсистемы МРАСУ являются комплексы программ №1 и №3,

осуществляющие соответственно мониторинг вторжений противника в функционирование системы, поиск уязвимостей в его средствах и генерирование информационных атак.

При вычислении значений параметров (198) используем следующий прием. Представим эти выражения как интегралы с переменным верхним пределом:

$$m_i(t) = \int_0^t p_i(\Theta) d\Theta, \quad \sigma_i^2(t) = \int_0^t p_i(\Theta)(1 - p_i(\Theta)) d\Theta \quad (199)$$

и вычислим производные по переменной t :

$$\dot{m}_i = p_i(t), \quad \dot{\sigma}_i^2 = p_i(t)(1 - p_i(t)), \quad i = (\overline{1,5}). \quad (200)$$

Рассматривая полученные выражения как дифференциальные уравнения для определения функций $m_i(t)$ и $\sigma_i(t)$, запишем для них начальные условия, которые с учетом (199) примут вид:

$$m_i(0) = 0, \quad \sigma_i^2(0) = 0, \quad i = (\overline{1,5}). \quad (201)$$

Тогда искомые значения $m_i(\tau)$ и $\sigma_i^2(\tau)$ получаются как результаты численного интегрирования [Вержбицкий В.М.] при $t = \tau$ расширенной системы уравнений (195), (200) с начальными условиями (196), (201).

Результаты расчета характеристик (198) с использованием предлагаемого приема приведены в Таб. 14.

Таблица 14

КП, i	1	2	3	4	5
$m_i(\tau), \text{с}$	62,968	40,979	104,582	74,419	17,052
$\sigma_i^2(\tau), \text{с}$	46,747	34,916	67,194	54,886	16,010
$(m_i \pm 3\sigma_i), \text{с}$	62,968 \pm 6,837	40,979 \pm 5,909	104,582 \pm 8,197	74,419 \pm 7,408	17,052 \pm 4,001

Из этой таблицы следует, что условие $\sum_{i=1}^5 m_i(\tau) = \tau$ выполняется и максимальные затраты времени работы на интервале

[0,300]с будет у комплекса программ №3, который осуществляет такую важнейшую функцию средств активной защиты как поиск уязвимостей в средствах нападения на рассматриваемую подсистему МРАСУ.

Самостоятельная работа.

Разработать общую имитационную модель функционирования средств активной защиты ИС.

Отметим, что к рассмотренным в предыдущих лекциях перспективным методам защиты данных в ИС можно отнести задачу оптимальной маскировки хранимой в системе информации, изложенной в работе [Моисеев В.С., Дятчин В.В., Тутубалин П.И.].

25. Расчетная модель оценки информационной живучести АСУ специального назначения

Для АСУ специального назначения (АСУ СН) [Моисеев В.С., Козар А.Н., Дятчин В.В.] важным является такой критерий эффективности, как живучесть системы.

В общем случае под живучестью понимается способность систем к сохранению своих основных функций (хотя бы с допустимой потерей качества их выполнения) при воздействии факторов внешней среды катастрофического характера (неблагоприятных условий эксплуатации).

Живучесть АСУ СН предлагается оценивать с помощью следующих компонент этого понятия:

- 1) боевая живучесть системы – возможность противостояния физическим воздействиям противника;
- 2) функциональная живучесть – способность системы оперативно восстанавливать свои свойства при отказах и неисправностях аппаратных программных средств АСУ СН;

3) информационная живучесть – возможность системы защищаться от информационных атак противника, а также оперативно устранять их последствия.

Отметим, что понятие *информационная живучесть* (ИЖ) систем в настоящее время является недостаточно разработанным.

При формализации этого показателя эффективности предлагается учитывать такие характеристики, как поток информационных атак, вероятность их пресечения, а также вероятность устранения успешных атак за заданное время.

Естественно потребовать, чтобы значение показателя ИЖ достигало на интервале времени функционирования системы своего максимума.

Важным компонентом этого критерия является достижение информационной безопасности (ИБ) разрабатываемой и эксплуатируемой АСУ СН. Характерной особенностью такой безопасности является тот факт, что её обеспечение должно осуществляться на конечном интервале времени каждого применения системы. Все это требует разработки специальных методов обеспечения ИЖ систем рассматриваемого класса.

Вопросам обеспечения информационной безопасности (ИБ) в различных сферах деятельности человека на сегодняшний день уделяется особое внимание. Об этом можно судить по возрастающему числу сообщений в средствах массовой информации, свидетельствующих об инцидентах связанным как с успешными, так и неуспешными попытками получить несанкционированный доступ к конфиденциальной информации. Степень важности сохранения таких данных значительно возрастает с уровнем их конфиденциальности, что влечет за собой необходимость решения актуальной и

злободневной задачи обеспечения ИБ систем их хранящих, использующих и обрабатывающих.

При этом одним из ключевых условий функционирования АСУ СН за счёт сферы их применения является обеспечение информационной безопасности. На фоне этого можно отметить отсутствие теоретических основ построения этого класса ИС и, главное, методов и средств обеспечения их ИБ.

Следует отметить, что само по себе понятие *информационной безопасности* определено на сегодня не однозначно и его зачастую сводят к другому, не менее вольно трактуемому, понятию *защиты информации* (ЗИ). С нашей точки зрения на практике следует применять понятия *информационной безопасности* и *защиты информации*, взяв за основу ГОСТ 51898-2002 и ГОСТ 50922-2006.

В соответствии с ними в прикладной информатике должны быть использованы следующие понятия *информационной безопасности* и *защиты информации*.

Информационная безопасность – допустимая, приемлемая в рассматриваемых условиях, вероятностная характеристика, определяющая возможный ущерб, который может быть нанесён информационной системе в результате несанкционированного доступа к её конфиденциальным данным, нарушения их целостности и доступности, а так же при нарушении целостности и доступности её программного обеспечения.

Защита информации – действия направленные на обеспечение ИБ системы (данных, программ, аппаратных средств и их возможных сочетаний).

Таким образом, опираясь на эти определения можно заключить, что понятие *информационной безопасности* по отношению к понятию *защита информации* не является конку-

рирующим, а по существу они являются объектами различной природы – первое определяет набор признаков и характеристик информационной системы, а второе определяет направление их достижения и обеспечения.

Отметим, что такой показатель как «Живучесть», особенно важен для военных АСУ СН. Для полевых АСУ СН предлагается использовать следующую совокупность основных критериев их эффективности: W_1 – производительность системы; W_2 – затраты времени на сопровождение и модификацию ПО системы и данных, хранящихся в системе; W_3 – живучесть; W_4 – затраты времени на развертывание и свертывание системы; W_5 – стоимость системы.

Относительно небольшое число критериев таких систем необходимо для удобства принятия проектных и эксплуатационных решений при достижении следующих целей:

$$\begin{aligned} W_1 &\rightarrow \max; & W_2 &\rightarrow \min; & W_3 &\rightarrow \max; \\ W_4 &\rightarrow \min; & W_5 &\rightarrow \min, \end{aligned}$$

что осуществляется с помощью реализации широко известных общетехнических решений, принимаемых на стадии проектирования АСУ, как систем специального (военного) назначения.

Максимизации показателя W_1 способствует применение в составе перспективных полевых АСУ (МРАСУ) распределенного информационно-вычислительного кластера (РИВК).

Минимизацию критерия W_2 в процессе эксплуатации систем предлагается осуществлять применением средств дистанционного обслуживания их программного обеспечения (ПО) и информационного обеспечения (ИО).

Характерной особенностью полевых АСУ СН (МРАСУ) является выполнение их функций на фиксированном интервале времени $[0, \tau]$. В этой связи предлагается рассматривать следующие функциональные критерии эффективности данного класса информационных систем (ИС):

$$W_1(\tau) \rightarrow \max; \quad W_2(\tau) \rightarrow \min; \quad W_3(\tau) \rightarrow \max.$$

При этом достижение первых двух целей, приведенных в этих выражениях, связано с обеспечением на каждом интервале $[0, \tau]$ времени функционирования системы ее максимальной живучести. Для других классов ИС, в частности для существующих к настоящему времени систем общего (гражданского) назначения величина $\tau = \infty$.

Рассмотрим состав функциональной характеристики $W_3(\tau)$. Будем считать, что она представляется вектором:

$$W_3(\tau) = (W_{31}(\tau), W_{32}(\tau), W_{33}(\tau)),$$

где $W_{31}(\tau)$ – показатель боевой живучести системы на интервале времени $[0; \tau]$; $W_{32}(\tau)$ – показатель функциональной живучести МРАСУ в этом интервале; $W_{33}(\tau)$ – показатель информационной живучести системы.

Цель достижения максимальной живучести системы можно декомпозировать в виде следующей совокупности локальных целей:

$$W_{31}(\tau) \rightarrow \max, \quad W_{32}(\tau) \rightarrow \max, \quad W_{33}(\tau) \rightarrow \max.$$

Для расчета и оптимизации значения $W_{31}(\tau)$ должны использоваться существующие мероприятия и методики. В качестве примеров таких решений можно указать на размещение элементов системы в бронированных образцах вооружений и военной техники (ВВТ), маскировку, в том числе радиомаскировку системы, а также ее охрану.

Для обеспечения максимальной боевой живучести в настоящее время практикуется оперативная смена мест размещения АСУ СН по истечению времени τ , которое для существующих систем составляет величину от 20 минут до 2 часов.

Показатель $W_{32}(\tau)$ связан с функциональной надежностью аппаратных и программных средств системы и возможность их быстрого восстановления на интервале $[0, \tau]$. Для расчета функциональной надежности и отказоустойчивости МРАСУ можно использовать методы, изложенные в монографии [Зиновьев П.А., Мейко А.В., Моисеев В.С.]. Достижению этой цели способствует применение в перспективных системах РИВК, как отказоустойчивого аппаратно-программного средства.

Отметим, что отказавшие в процессе эксплуатации системы компоненты ее ПО восстанавливаются дистанционно.

Перейдем к формированию показателя $W_{33}(\tau)$. Следует отметить, что в существующих работах по живучести ИС [Додонов А.Г., Ландэ Д.В.] отсутствуют конкретные методики расчета показателя их информационной живучести, которые можно было бы применять при разработке и эксплуатации систем.

Показатель информационной живучести (ИЖ) системы в общем виде предлагается оценивать с помощью следующего выражения:

$$W_{33}(\tau) = 1 - \sum_{k=1}^K P_A(\tau, \lambda, k) P_{\text{НИБ}}(k) P_B(\tau, \Theta). \quad (202)$$

Здесь $P_A(\tau, \lambda, k)$ – безусловная вероятность того, что на интервале времени $[0, \tau]$ функционирования МРАСУ противник осуществит k информационных атак с интенсивностью λ ,

$k = (\overline{1, K})$, где K – максимальное число информационных атак на рассматриваемую систему в интервале времени ее работы; $P_{\text{НИБ}}(k)$ – условная вероятность хотя бы одного нарушения ИБ системы по результатам k атак; $P_B(\tau, \Theta)$ – условная вероятность того, что случайные затраты времени Θ на ликвидацию последствий атак превысят допустимое время $\nu_{\text{доп}}$.

Для проведения практических расчетов ИЖ конкретизируем формулу (202) при следующих инженерных предположениях:

- 1) поток информационных атак на систему считается пуассоновским потоком с постоянной интенсивностью λ ;
- 2) затраты времени Θ распределены по показательному закону с функцией распределения $F(\theta) = P(\Theta < \theta) = 1 - e^{-\alpha\theta}$;
- 3) величина $\nu_{\text{доп}}$ рассматривается как допустимая доля времени τ функционирования системы, т.е. $\nu_{\text{доп}} = \alpha \cdot \tau$, где $\alpha \in (\overline{0, 1})$ – заданная величина.

Согласно этим предположениям будем использовать следующие зависимости:

$$\begin{aligned} P_A(\tau, \lambda, k) &= \left(\frac{\lambda\tau}{k!} \right)^k e^{-\lambda\tau}; \\ P_{\text{НИБ}}(k) &= 1 - (1 - p_{\text{НИБ}})^k; \\ P_B(\tau, \Theta) &= P(\Theta > \alpha\tau) = e^{-\alpha\tau}, \end{aligned} \tag{203}$$

где $p_{\text{НИБ}}$ – допустимое значение вероятности нарушения ИБ системы.

Подставляя (203) в (202), получим после несложных преобразований выражение вида:

$$W_{33}(\tau) = 1 - e^{-\tau(\lambda+\alpha)} \sum_{k=1}^K \frac{(\lambda\tau)^k}{k!} \left(1 - (1 - p_{\text{НИБ}})^k\right). \quad (204)$$

Проведем исследование полученной расчетной модели ИЖ.

Максимум $\bar{W}_{33}(\tau)$ достигает при $p_{\text{НИБ}} = 0$, т.е. при наличии в АСУ СН абсолютной защиты от информационных атак противника. При $p_{\text{НИБ}} = 1$ (отсутствии защиты) показатель ИЖ $W_{33}(\tau)$ достигает минимального значения:

$$\hat{W}_{33}(\tau) = 1 - e^{-\tau(\lambda+\alpha)} \sum_{k=1}^K \frac{(\lambda\tau)^k}{k!}. \quad (205)$$

Проверим на интервале $[0, \tau]$ наличие экстремумов показателя ИЖ. Для этого введем текущее время t функционирования АСУ СН и перепишем выражение (204) в следующем виде:

$$W_{33}(t) = 1 - e^{-t(\lambda+\alpha)} \sum_{k=1}^K \frac{(\lambda t)^k}{k!} \left(1 - (1 - p_{\text{НИБ}})^k\right), \quad t \in [0, \tau]. \quad (206)$$

Используя необходимое условие экстремума функции (206), запишем следующее уравнение для выделения ее стационарных точек:

$$\begin{aligned} \frac{dW_{33}}{dt} &= (\lambda + \alpha) e^{-t(\lambda+\alpha)} \sum_{k=1}^K \frac{(\lambda t)^k}{k!} \left(1 - (1 - p_{\text{НИБ}})^k\right) - \\ &- e^{-t(\lambda+\alpha)} \sum_{k=1}^K \frac{k\lambda^k t^{k-1}}{k!} \left(1 - (1 - p_{\text{НИБ}})^k\right) = 0. \end{aligned}$$

Так как $e^{-t(\lambda+\alpha)} \neq 0$, то получим уравнение вида:

$$\sum_{k=1}^K \left((\lambda + \alpha) \frac{(\lambda t)^k}{k!} - \frac{\lambda^k t^{k-1}}{(k-1)!} \right) \left(1 - (1 - p_{\text{НИБ}})^k\right) = 0.$$

Это уравнение представляет собой алгебраическое уравнение K -го порядка, среди корней t_1, t_2, \dots, t_k которого могут быть

действительные корни, принадлежащие интервалу $[0, \tau]$. Это означает, что на интервале $[0, \tau]$ показатель ИЖ (206) может принимать максимальные и (или) минимальные значения.

Так как время работы системы τ мало, то определение этих моментов можно осуществить графическим методом.

При реальных значениях вероятности $p_{\text{НИБ}} > 0$ показатель ИЖ лежит в пределах:

$$\hat{W}_{33}(\tau) < W_{33}(\tau) < 1. \quad (207)$$

Для конкретной АСУ СН (МРАСУ) оптимальное значение показателя ИЖ достигается обоснованным выбором значения $p_{\text{НИБ}}$ или вероятности $p_{\text{ИБ}} = 1 - p_{\text{НИБ}}$ обеспечения ИБ системы, а также мероприятиями по оптимальному выбору величин T_B и α , которые оценивают затраты времени на ликвидацию последствий атак противника и долю времени возможной остановки функционирования системы.

Пример.

Пусть некоторая МРАСУ должна работать на каждом месте ее размещения в течение 20 минут.

Предполагается, что противник осуществляет $K = 8$ информационных атак с интенсивностью $\lambda = 0,5$ мин. Для успешного выполнения функций рассматриваемой системы при ее разработке и эксплуатации были выбраны следующие значения параметров обеспечения ИБ и ИЖ:

$$p_{\text{НИБ}} = 0,02; \quad T_B = 0,5 \text{ мин}; \quad \alpha = 0,05.$$

Последний параметр указывает на то, что на ликвидацию последствий каждой из пропущенных средствами обеспечения ИБ системы атак должно отводиться не более 5% времени функционирования системы, т.е. $\nu_{\text{доп}} = 1$ мин.

Подставляя эти исходные данные в выражения (204) и (205), получим, что:

$$W_{33}(20) = 0,994 \text{ и } \hat{W}_{33}(20) = 0,955.$$

Из Рис. 77 следует, что при $p_{\text{НИБ}} = 0,02$ минимум функции (206) достигается в точке $t = 7,7$ мин со значением $W_{33}(7,7) = 0,967$. При $p_{\text{НИБ}} = 1$ минимум функции (206) достигается в точке $\tau = 3,7$ мин со значением $\hat{W}_{33}(3,7) = 0,418$ (Рис. 78).

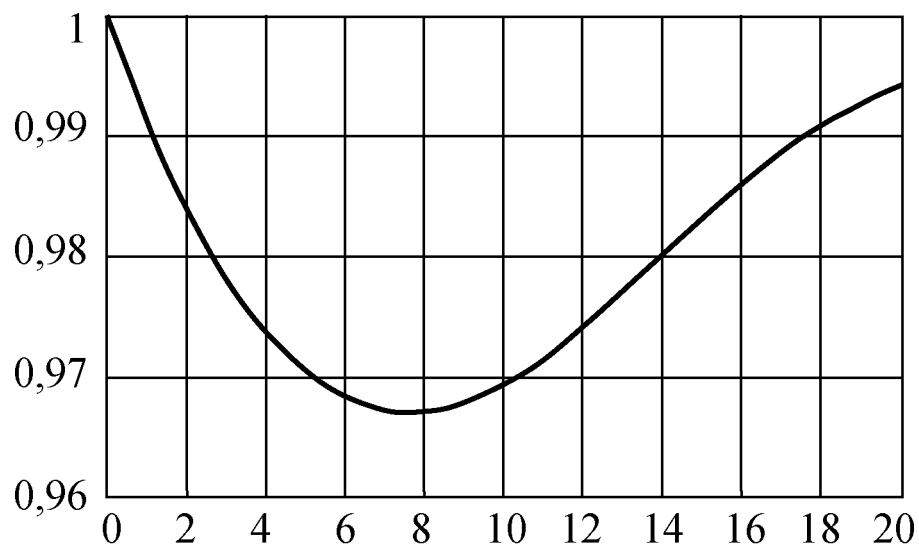


Рис. 77

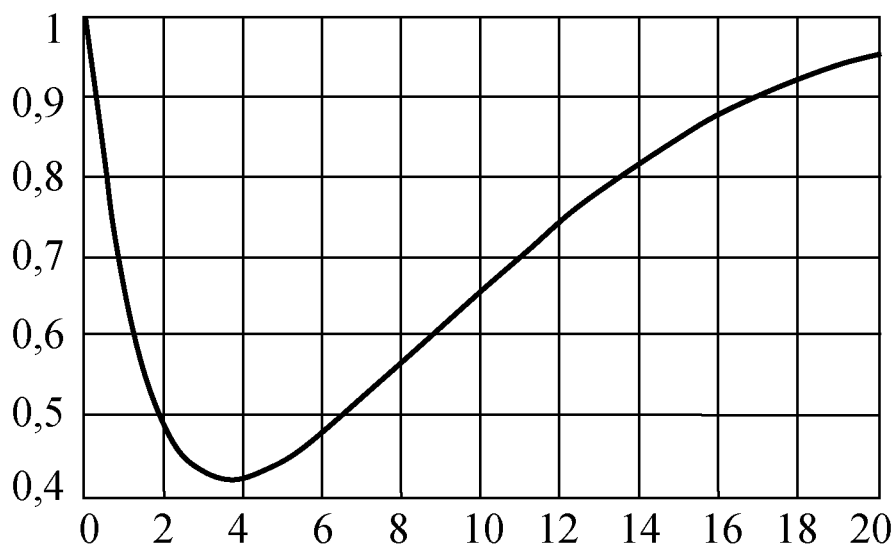


Рис. 78

Установлено, что показатель ИЖ системы убывает с уменьшением допустимого значения вероятности $p_{\text{НИБ}}$ нарушения ИБ системы и с увеличением затрат времени на ликвидацию последствий атак, характеризуемых показателем T_B . При этом увеличение значения T_B вносит более существенный вклад в снижение показателя ИЖ системы по сравнению с уменьшением значения $p_{\text{НИБ}}$ (вплоть до единицы).

Значение ИЖ увеличивается с ростом величины α , определяющей допустимую долю времени, затрачиваемой на устранение последствий информационных атак.

В заключение лекции необходимо отметить, что при решении проблемы информационной безопасности систем различного назначения необходимо опираться не на интуитивные подходы, а на конкретные модели и методы прикладной математики.

ЗАКЛЮЧЕНИЕ

В предлагаемой книге рассмотрены простейшие математические модели задач прикладной информатики.

Автор надеется на активное и систематизированное развитие математических основ прикладной информатики, чтобы ее как науку не постигла печальная участь кибернетики.

ИСПОЛЬЗОВАННАЯ И РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. *Амосов Н.М.* Алгоритмы разума. – К.: Наукова думка, 1979. – 225 с.
2. *Анохин И.А.* Развитие вычислительной техники и систем на ее основе // Новости искусственного интеллекта, 1994, №1. С. 26-69.
3. *Бройдо В.Л.* Вычислительные системы, сети и телекоммуникации. – СПб: Питер, 2003. – 456 с.
4. *Бутузова А.В., Моисеев В.С., Тутубалин П.И.* Теоретические основы информатизации службы скорой медицинской помощи. – Казань: РИЦ «Школа», 2011. – 242 с. (Серия «Современная прикладная математика и информатика»).
5. Введение в математическое моделирование / Под ред. П.В. Трусова. – М.: «Интернет Инжиниринг», 2000. – 336 с.
6. *Вентцель Е.С.* Исследование операций. – М.: Сов. радио, 1972. – 552 с.
7. *Вержбицкий В.М.* Основы численных методов. – Высшая школа, 2002. – 840 с.
8. *Винер Н.* Кибернетика или управление и связь в животном и машине. – М.: Сов. радио, 1958. – 232 с.
9. *Винер Н.* Новые главы кибернетики. – М.: Сов. радио, 1963. – 208 с.
10. *Винер Н.* Я – математик. – М.: Наука, 1967. – 356 с.
11. *Герасименко А.В.* Проблемы защиты информации в системах их обработки // Зарубежная радиоэлектроника, №2, 1989. – С. 5-21.
12. *Гинатуллин И.А., Зиновьев П.А., Моисеев В.С.* Модель для исследования динамики изменений количества пользователей в корпоративных информационных системах // Иссле-

дования по информатике. Выпуск 6. – Казань: Отечество, 2003. – С. 111-120.

13. *Гинатуллин И.А., Моисеев В.С.* Двухуровневая математическая модель корпоративных информационных систем на основе линейных стохастических сетей // Исследования по информатике. Выпуск 7. – Казань: Отечество, 2004. – С. 89-100.

14. *Глушков В.М., Валах В.Я.* Что такое ОГАС? – М.: Наука, 1981. – 160 с.

15. *Глушков В.М.* Введение в АСУ. – К.: Техніка, 1974. – 317 с.

16. *Глушков В.М.* Макроэкономические модели и принципы построения ОГАС. – М.: Статистика, 1975. – 160 с.

17. *Глушков В.М.* Основы безбумажной информатики. – М.: Наука, 1987. – 552 с.

18. *Глушков В.М. и др.* Сети ЭВМ. – М.: Радио и связь, 1977. – 279 с.

19. *Гнеденко Б.В.* Курс теории вероятностей. – М.: Наука, 1988. – 488 с.

20. *Данилевский Ю.Г., Петухов И.А., Шибанов В.С.* Информационная технология в промышленности – Л.: Машиностроение, 1988. – 283 с.

21. *Диго С.М.* Базы данных: проектирование и использование. – М.: Финансы и статистика, 2005. – 592 с.

22. *Доропов А.Г., Ландэ Д.В.* Живучесть информационных систем. – К.: Наукова думка, 2011. – 256 с.

23. *Ершов А.П.* Введение в теоретическое программирование. – М.: Наука, 1977. – 288 с.

24. *Жарова О.Ю., Федорова В.А.* Метод определения типа атаки по статистическим параметрам сетевого трафика // Вопросы радиоэлектроники, №2, 2016. – С. 39-43.

25. *Зайченко Ю.П., Гонта Ю.В.* Структурная оптимизация сетей ЭВМ. – К. Техніка, 1986. – 167 с.

26. *Залецанский Б.Д., Чернихов Д.Я.* Кластерная технология и живучесть глобальных автоматизированных систем. – М.: Финансы и статистика, 2005. – 384 с.

27. *Зарубин В.С.* Математическое моделирование в технике. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. – 496 с.

28. *Зиновьев П.А., Мейко А.В., Моисеев В.С.* Инженерные методы расчета функциональной надежности и живучести корпоративных информационных систем. – Казань: Отечество, 2009. – 256 с. (Серия «Современная прикладная математика и информатика»).

29. *Зиновьев П.А., Моисеев В.С., Фазылзянов Ф.М.* К теории региональной информатизации. // Исследования по информатике. Вып. 7. – Казань: Отечество, 2004. – С. 3-20.

30. *Ивани Л., Смилянский Р.П.* Элементы теоретического программирования. – М.: Изд. МГУ, 1985.

31. Информатика. Изд. 3-е перераб., под ред. Н.В. Макаровой. – М.: Финансовая статистика, 1999. – 768 с.

32. Информатика. Энциклопедический словарь, под ред. Д.А. Поспелова. – М.: Педагогика Пресс, 1994. – 352 с.

33. *Колмогоров А.Н., Драгалин А.Г.* Математическая логика. – М.: Изд. МГУ, 1984. – 119 с.

34. *Колмогоров А.Н., Фомин С.В.* Элементы теории множеств и функционального анализа. – М.: Физматлит, 2006. – 572 с.

35. Компьютер и задачи выбора. – М.: Наука, 1989. – 208 с.

36. *Кристофидес Н.* Теория графов. Алгоритмический подход – М.: Мир, 1978. – 432 с.

37. *Кузин Л.Т.* Основы кибернетики. В 2-х томах. Т. 2. Основы кибернетических моделей. – М.: Энергия, 1979. – 584 с.
38. *Кучин Б.Л., Якушенко Е.В.* Управление развитием экономических систем. – М.: Экономика, 1990. – 256 с.
39. *Куратовский К., Мостовский А.* Теория множеств. – М.: Мир, 1970. – 416 с.
40. *Козар А.Н., Моисеев В.С.* Информационные технологии оптимального применения управляемых артиллерийских сарядов: монография. – Казань: РЦ МКО, 2012. – 348 с. (Серия «Современная прикладная математика и информатика»).
41. *Котов В.Е.* Сети Петри. – М.: Наука, 1984. – 160 с.
42. *Кофман А., Крюон Р.* Массовое обслуживание. Теория и приложения. – М.: Мир, 1965. – 302 с.
43. *Кузнецов Н.А., Кульба В.В., Микрин В.А.* Информационная безопасность систем организационного управления. Теоретические основы. В 2-х т. Ин-т проблем передачи информации РАН. – М.: Наука, 2006.
44. *Кук Д., Бейз Г.* Компьютерная математика. – М.: Наука, 1990. – 385 с.
45. *Куликовский Л.Ф., Мотов В.В.* Теоретические основы информационных процессов. – М.: Высшая школа, 1987. – 248 с.
46. *Лукацкий А.В.* Обнаружение атак. – СПб.: БХВ-Петербург, 2001. – 624 с.
47. *Мамиконов А.Г., Кульба В.В., Шелков А.Б.* Достоверность, защита и резервирование информации в АСУ. – М.: Энергоатомиздат, 1986. – 304 с. (Серия «Применение вычислительных машин в исследованиях и управлении производством»).
48. *Мартино Д.Х.* Технологическое прогнозирование. – М.: Прогресс, 1997. – 320 с.

49. Математическое моделирование / Под ред. Дж. Эндрюса и Р. Мак-Лоуна. М.: Мир, 1979. – 280 с.

50. Месарович М., Мако Д., Такахара И. Теория иерархических многоуровневых систем. – М.: Мир, 1973. – 340 с.

51. Миролюбов А.А., Солдатов М.А. Линейные однородные разностные уравнения. – М.: Наука, 1981. – 225 с.

52. Моисеев В.С., Дятчин В.В., Тутубалин П.И. К задаче маскировки конфиденциальных данных автоматизированных систем // Вестник КГТУ им. А.Н. Туполева, №2, 2007. – С. 55-58.

53. Моисеев В.С., Козар А.Н., Дятчин В.В. Информационная безопасность автоматизированных систем управления специального назначения. – Казань: КВАКИУ (ВИ), 2006. – 382 с.

54. Моисеев В.С., Николаев А.А. Анализ характеристик программных изделий на основе марковских динамических моделей // Вестник КГТУ им. А.Н. Туполева, №3, 2000. – С. 59-65.

55. Моисеев В.С., Тутубалин П.И. Вероятностная динамическая модель функционирования программных средств активной защиты мобильных распределенных АСУ // Информационные технологии, №6. – М.: Новые технологии, 2013. – С. 37-42.

56. Моисеев В.С., Тутубалин П.И. Об одном подходе к обеспечению активной защиты информационных систем // Вестник КГТУ им. А.Н. Туполева, №2, 2011. – С. 124-135.

57. Подиновский В.В., Ногин В.Д. Паретооптимальные решения многокритериальных задач. – М.: Наука, 1982. – 255 с.

58. Поспелов Д.А. Становление информатики в России. – В сб. Очерки истории информатики в России // Составители –

Д.А. Поспелов, Я.И. Фет. – Новосибирск, Научно-издательский центр ОИГГМ, 1998. – С. 7-44.

59. *Раскин Л.Т.* Анализ сложных систем и элементы теории оптимального управления. – М.: Сов. радио, 1976. – 344 с.

60. *Самарский А.А., Михайлов А.П.* Математическое моделирование. – М.: Физматлит, 2002. – 320 с.

61. *Сигал И.Х., Иванова А.В.* Введение в прикладное дискретное программирование. – М.: Физматлит, 2002. – 256 с.

62. *Скембрей Дж., Мак-Клар Ст., Куриц Дж.* Секреты хакеров. Безопасность сетей – готовые решения. – М.: Вильямс, 2001. – 656 с.

63. *Таха Х.* Введение в исследование операций: В 2-х книгах. Кн. 2. – М.: Мир, 1985. – 496 с.

64. *Урсул А.Д.* Становление информационного общества и переход к устойчивому развитию // Проблемы информатики, №4, 1997. – С. 13-20.

65. *Харари Ф.* Теория графов. – Едиториал УРСС, 2003. – 296 с.

66. *Цегелик Г.Г.* Системы распределенных баз данных. – Львов: Свит, 1990. – 168 с.

67. *Шеннон Р.* Имитационное моделирование систем – искусство и наука. – М.: Мир, 1978. – 420 с.

68. *Эшби У.Р.* Введение в кибернетику. – М.: Изд. иностр. литер., 1959. – 144 с.

69. *Юсупов Р.М., Заболотский В.П.* Научно-методологические основы информатизации. – СПб.: Наука, 2000. – 455 с.

70. *Янбых Г.Ф., Эттингер Б.Я.* Методы анализа и синтеза сетей ЭВМ. – Л.: Энергия, 1980.

Виктор Сергеевич Моисеев

**ЛЕКЦИИ ПО МАТЕМАТИЧЕСКОМУ
МОДЕЛИРОВАНИЮ В ПРИКЛАДНОЙ
ИНФОРМАТИКЕ**

Научно-учебное издание

Подписано к печати 27.04.2018.

Формат 60х84^{1/16}. Бумага офсетная.

Гарнитура «Times». Печать ризографическая.

Усл. печ. 17,90 л. Печ. 19,25 л. Тираж 100 экз. Заказ № 62.

420111, Казань, Дзержинского, 9/1. Тел. сот.: +7(917)264-84-83.

Отпечатано с готового оригинал-макета
в редакционно-издательском центре «Школа»

Е-mail: ric-school@yandex.ru